



**Universidade do Minho**  
Escola de Engenharia

Ana Margarida Aires Cancelinha Oliveira

**Monitorização de máquinas industriais  
recorrendo a técnicas de Data Mining e  
Machine Learning**

Dissertação de Mestrado

Mestrado Integrado em Engenharia Eletrónica Industrial  
e Computadores

Trabalho efetuado sob a orientação de  
Professor Doutor Sérgio Adriano Fernando Lopes  
Doutora Carla Susana Correia Pereira

Mês [de submissão da dissertação] de 201X

## DECLARAÇÃO

Nome: Ana Margarida Aires Cancelinha Oliveira

Endereço eletrónico: anairesoli@gmail.com Telefone: 938057067

Bilhete de Identidade/Cartão do Cidadão: 14418763

Título da dissertação: Monitorização de máquinas industriais recorrendo a técnicas de Data Mining e Machine Learning

Orientador/a/es:

Professor Doutor Sérgio Adriano Fernando Lopes

Doutora Carla Susana Correia Pereira

Ano de conclusão: 2017

Mestrado em Engenharia Eletrónica Industrial e Computadores

DE ACORDO COM A LEGISLAÇÃO EM VIGOR, NÃO É PERMITIDA A REPRODUÇÃO DE QUALQUER PARTE DESTA TESE/TRABALHO.

Universidade do Minho, \_\_\_\_/\_\_\_\_/\_\_\_\_

Assinatura:



## AGRADECIMENTOS

Embora a realização de uma dissertação seja, de certa forma, um processo solitário, existem sempre pessoas que contribuem, direta ou indiretamente, para seu desenvolvimento. Assim, quero agradecer a todas as pessoas que me ajudaram, com esta tese de mestrado, a finalizar uma etapa muito importante na minha vida.

Começo por agradecer aos meus pais, Margarida e João, que me deram toda a motivação, amor e apoio inestimável. Gostaria de lhes agradecer a oportunidade de ter frequentado um excelente curso numa grande instituição.

Quero agradecer ao meu irmão João Pedro, por estar sempre presente com palavras e gestos de coragem que me ajudaram nesta etapa.

Um agradecimento muito especial ao Júlio Alves, pelo amor, carinho, paciência e presença. Obrigada por me ajudares a crescer e a ser resiliente.

Agradeço ao Professor Doutor Sérgio Lopes e Doutora Carla Pereira (orientadores) pela disponibilidade, ajuda, dedicação e orientação ao longo desta dissertação.

Quero agradecer à Controlar a confiança depositada em mim para o desenvolvimento desta dissertação. Um agradecimento adicional ao departamento IDI, ao Pedro e Rafael, pelo companheirismo e boa disposição.

Um muito obrigado a todos os meus amigos de infância, Kika, Zé, Rita, Salomé e Bárbara e aos amigos ganhos ao longo deste percurso, os PATOS DA SHEET!!!!, por todo o apoio, força e camaradagem.

Agradeço a todos os que me ajudaram a ser eu, que acreditaram em mim e me apoiaram incondicionalmente, levo-vos no meu coração.

Obrigada.



## RESUMO

Hoje em dia, está-se perante uma 4ª revolução industrial marcada pela utilização das tecnologias da informação nos processos industriais.

A monitorização contínua de vários parâmetros que contribuem para o bom funcionamento das máquinas é um processo que engloba uma grande recolha de dados e a sua análise. Estes dados recolhidos contêm muita informação útil que pode ser utilizada para uma melhor execução de tarefas ou tomada de decisões.

Um sistema de monitorização, que recorre a *Data Mining* e *Machine Learning*, concede grandes vantagens às empresas, como a redução de custos de manutenção, tempo de inatividade, entre outras.

Nesta dissertação, desenvolveu-se um sistema de deteção e previsão de falha de componentes rotativos, aplicando técnicas de *data mining* e *machine learning* a dados de vibração extraídos de máquinas industriais.

Numa primeira fase, os dados de vibração armazenados foram pré-processados com diferentes métodos, de forma a extrair as características mais relevantes para o problema em causa. De seguida, foram aplicados vários algoritmos de *Data Mining* com o objetivo de obter toda a informação útil. Após a obtenção dessa informação, aplicaram-se técnicas de *Machine Learning* com o objetivo de tomar as melhores decisões sobre a máquina, considerando a informação extraída.

Vários casos foram estudados e implementados, contudo os melhores resultados surgiram da utilização das técnicas de pré-processamento *Discrete Wavelet Transform* (DWT) e *Empirical Mode Decomposition* (EMD) associadas ao classificador *Support Vector Data Description* (SVDD).

**Palavras-Chave:** Monitorização de máquinas, sinais de vibração, Noise, Vibration and Harshness (NHV), *Data Mining*, *Machine Learning*



## ABSTRACT

Today, we are facing a 4th industrial revolution marked by the use of information technology in industrial processes.

The continuous monitoring of several parameters which ensure proper operation of a machine, it's a process that includes massive data collection and analysis. The collected data has useful information that can be used for a better task execution or decision making.

A monitoring system, which uses Data Mining and Machine Learning, brings great advantages to industry like the reduction of maintenance costs and downtime.

In this dissertation, a failure prediction and detection system for rotating components is developed by applying Data Mining and Machine Learning techniques to vibration data extracted from industrial machines.

Initially, the storage vibration data is preprocessed with different methods, in order to extract the most relevant features for the problem in question. Next, several Data Mining algorithms were applied to obtain all useful information. After obtaining the information, Machine Learning techniques were applied to make the best decisions about the machine, considering the information extracted.

Several cases were studied and implemented, but the best results were obtained using the Discrete Wavelet Transform (DWT) and Empirical Mode Decomposition (EMD) preprocessing techniques associated with the Support Vector Data Description (SVDD) classifier.

**Keywords:** Machine Monitoring, Vibration Signals, Noise, Vibration and Harshness (NHV), Data Mining, Machine Learning



# ÍNDICE

|                                                                         |    |
|-------------------------------------------------------------------------|----|
| 1. Introdução.....                                                      | 1  |
| 1.1 Motivação .....                                                     | 2  |
| 1.2 Objetivos.....                                                      | 3  |
| 1.3 Estrutura do documento .....                                        | 4  |
| 2. Estado da Arte.....                                                  | 5  |
| 2.1 Pré-processamento .....                                             | 5  |
| 2.1.1 Domínio do tempo .....                                            | 6  |
| 2.1.2 Domínio da frequência.....                                        | 7  |
| 2.1.3 Domínio do tempo-frequência.....                                  | 9  |
| 2.2 Extração de conhecimento .....                                      | 21 |
| 2.2.1 Algoritmos supervisionados .....                                  | 21 |
| 2.2.2 Aprendizagem não supervisionada .....                             | 29 |
| 3. Metodologia.....                                                     | 33 |
| 3.1 Dados .....                                                         | 33 |
| 3.1.1 Dados de referência .....                                         | 33 |
| 3.1.2 Dados disponibilizados pela empresa de acolhimento.....           | 34 |
| 3.1.3 Dados que caracterizam falhas específicas.....                    | 36 |
| 3.2 Ferramentas utilizadas .....                                        | 37 |
| 3.2.1 Matlab .....                                                      | 37 |
| 3.3 Abordagem e Métodos.....                                            | 38 |
| 3.3.1 Filtragem .....                                                   | 39 |
| 3.3.2 Pré-processamento e Extração de características .....             | 39 |
| 3.3.3 Classificação e Aprendizagem do Sistema .....                     | 41 |
| 4. Abordagem estatística .....                                          | 43 |
| 4.1 Blind Source Separation + Curtograma Rápido e análise envelope..... | 43 |
| 4.1.1 Blind Source Separation .....                                     | 43 |

|       |                                                                          |    |
|-------|--------------------------------------------------------------------------|----|
| 4.1.2 | Curtograma rápido e análise envelope .....                               | 45 |
| 4.2   | Wavelet Denoise + Wavelet Packet Transform .....                         | 48 |
| 4.2.1 | Wavelet Denoise .....                                                    | 49 |
| 4.2.2 | Wavelet Packet Transform .....                                           | 50 |
| 4.3   | Conclusão.....                                                           | 54 |
| 5.    | Abordagem de aprendizagem supervisionada do sistema.....                 | 55 |
| 5.1   | Discrete Wavelet Transform + Support Vector Machine .....                | 55 |
| 5.1.1 | Discrete Wavelet Transform .....                                         | 56 |
| 5.1.2 | Support Vector Machine .....                                             | 56 |
| 5.2   | Discrete Wavelet Transform + Fuzzy C-Means Clustering .....              | 60 |
| 5.2.1 | Discrete Wavelet Transform .....                                         | 61 |
| 5.2.2 | Fuzzy C-Means Clustering .....                                           | 61 |
| 5.3   | Conclusão.....                                                           | 63 |
| 6.    | Abordagem de aprendizagem não supervisionada .....                       | 65 |
| 6.1   | Discrete Wavelet Transform/Empirical Mode Decomposition .....            | 65 |
| 6.2   | One-Class Support Vector Machine e Support Vector Data Description ..... | 66 |
| 6.2.1 | DWT/EMD + OCSVM .....                                                    | 68 |
| 6.2.2 | DWT/EMD + SVDD.....                                                      | 74 |
| 6.3   | Conclusão.....                                                           | 83 |
| 7.    | Conclusões e trabalho futuro .....                                       | 84 |
| 7.1   | Conclusão.....                                                           | 84 |
| 7.2   | Trabalho futuro.....                                                     | 86 |
|       | Bibliografia .....                                                       | 88 |

## LISTA DE FIGURAS

|                                                                                                                         |    |
|-------------------------------------------------------------------------------------------------------------------------|----|
| Figura 1-Sinais de Vibração de caixas de engrenagens em boas e más condições (Baqqar et al., 2011).....                 | 6  |
| Figura 2-Aplicação da FFT (George Dallas, 2014) .....                                                                   | 8  |
| Figura 3-Frame ou Time Window utilizado no domínio da frequência (imagem adaptada de (Grosse & Reinhardt, 2002)).....   | 8  |
| Figura 4-Método STFT (adaptado de (Gutierrez-Osuna, 2002)) .....                                                        | 10 |
| Figura 5-Tipos de Wavelets (George Dallas, 2014) .....                                                                  | 11 |
| Figura 6-Representação tempo e escala resultante da aplicação da transformada de wavelets (Haar, 1998).....             | 12 |
| Figura 7-Wavelet com diferentes escalas (imagem adaptada de (George Dallas, 2014)).....                                 | 12 |
| Figura 8-Resultado final do vetor x, depois de aplicar uma transformada discreta wavelet de nível 3 (Graps, 2004) ..... | 15 |
| Figura 9-Procedimento realizado a cada nível da banda de frequência do método DWT (Graps, 2004).....                    | 15 |
| Figura 10-Decomposição em árvore da WTP de nível 3 .....                                                                | 16 |
| Figura 11-Fluxograma do método EMD (Yan et al., 2006).....                                                              | 18 |
| Figura 12-Envelope LMI (imagem adaptada de (Fang & Luo, 2009)).....                                                     | 19 |
| Figura 13-Método BSS (adaptado de (Han, Rafii, & Pardo, 2017)) .....                                                    | 20 |
| Figura 14-Camadas de uma rede neuronal .....                                                                            | 22 |
| Figura 15-Classificador SVM.....                                                                                        | 24 |
| Figura 16-Comparação do sistema de classificação clássico com a lógica difusa.....                                      | 26 |
| Figura 17-Exemplo do processo de decisão do algoritmo KNN .....                                                         | 27 |
| Figura 18- Árvore de decisão para o problema de jogar Ténis consoante o tempo.....                                      | 28 |
| Figura 19-Método de classificação OCSVM .....                                                                           | 30 |
| Figura 20-Método de classificação da SVDD.....                                                                          | 31 |
| Figura 21-Plataforma de teste e esquema de posicionamento dos acelerómetros (Qiu et al., 2006).....                     | 34 |
| Figura 22-Posicionamento do sensor (D) na máquina teste.....                                                            | 35 |

|                                                                                                                            |    |
|----------------------------------------------------------------------------------------------------------------------------|----|
| Figura 23-Exemplos de amostras de vibração que caracterizam o mau funcionamento do equipamento .....                       | 35 |
| Figura 24-Exemplos de amostras de vibração que caracterizam o bom funcionamento do equipamento .....                       | 36 |
| Figura 25-Motor elétrico utilizado na aquisição de dados de vibração (University, 2017) .....                              | 37 |
| Figura 26-Diagrama da implementação BSS + Curtograma rápido e análise envelope .....                                       | 43 |
| Figura 27-Diagrama da implementação 1 .....                                                                                | 43 |
| Figura 28-Sinal de vibração filtrado, utilizando o algoritmo Fast-ICA.....                                                 | 44 |
| Figura 29-Rápido curtograma do sinal de vibração sem falha (A) e com falha (B).....                                        | 45 |
| Figura 30-Árvore curtograma. Adaptado de [3].....                                                                          | 46 |
| Figura 31-Análise envelope dos sinais sem falha (A) e com falha (B).....                                                   | 47 |
| Figura 32-Diagrama da implementação WD + WPT.....                                                                          | 49 |
| Figura 33-Denoise do sinal de vibração .....                                                                               | 50 |
| Figura 34-Energia dos coeficientes WPT de amostras sem falha (A-dados de referência, B-dados com falhas específicas) ..... | 51 |
| Figura 35-Energia dos coeficientes WPT de amostras com falha (A-dados de referência, B-dados com falhas específicas) ..... | 52 |
| Figura 36-Histórico dos coeficientes de energia do conjunto de amostras de referência .....                                | 53 |
| Figura 37-Diagrama de implementação DWT + SVM.....                                                                         | 55 |
| Figura 38-Procedimento de contagem de amostras com falha para os dois testes realizados .....                              | 57 |
| Figura 39-Classificação da SVM com conjunto de total de amostras (conjunto de treino + conjunto de teste) .....            | 58 |
| Figura 40-Classificação da SVM com conjunto de teste .....                                                                 | 58 |
| Figura 41-Alguns resultados do vetor de probabilidades da classificação da SVM sem e com conjunto de treino.....           | 58 |
| Figura 42-Quantidade de amostras classificadas como falha nos gráficos da fig. 39 e 40 respectivamente .....               | 59 |
| Figura 43- Classificação da SVM com conjunto de total de amostras (conjunto de treino + conjunto de teste) .....           | 60 |
| Figura 44-Classificação da SVM com conjunto de teste .....                                                                 | 60 |
| Figura 45-Diagrama de implementação DWT + FCM.....                                                                         | 61 |

|                                                                                                   |                                     |
|---------------------------------------------------------------------------------------------------|-------------------------------------|
| Figura 46-Classificação FCM para: A) 100 amostras e B) 1000 amostras .....                        | 62                                  |
| Figura 47-Grau de associação de cada amostra aos dois clusters .....                              | 63                                  |
| Figura 48-Representação das características extraídas numa matriz sparse .....                    | 67                                  |
| Figura 49-Diagrama de implementação DWT + FCM.....                                                | 68                                  |
| Figura 50-Classificação OCSVM utilizando 8 energias extraídas pela DWT .....                      | 69                                  |
| Figura 51-Evolução da classe normal e de falha na classificação da fig. 50 .....                  | 70                                  |
| Figura 52-Classificação OCSVM com energias das 8 IMFs .....                                       | 71                                  |
| Figura 53-Evolução da classe normal e de falha na classificação da fig. 53 .....                  | 71                                  |
| Figura 54-Classificação com energias das 2 primeiras IMFs .....                                   | 71                                  |
| Figura 55-Evolução da classe normal e de falha na classificação da fig. 50 .....                  | 72                                  |
| Figura 56-Evolução da classe -1, em cada eixo, para cada método de pré-processamento...           | 72                                  |
| Figura 57-Classificação dos dados da Controlar utilizando a DWT .....                             | 73                                  |
| Figura 58-Classificação dos dados da Controlar utilizando a EMD .....                             | 74                                  |
| Figura 59-Procedimento de teste do conjunto de dados de referência .....                          | 76                                  |
| Figura 60-Procedimento de teste das amostras da Controlar .....                                   | 76                                  |
| Figura 61-Procedimento de teste das amostras da Controlar .....                                   | <b>Erro! Marcador não definido.</b> |
| Figura 62-Procedimento de teste do conjunto de dados de referência.....                           | <b>Erro! Marcador não definido.</b> |
| Figura 63-Diagrama de implementação DWT/EMD + SVDD.....                                           | <b>Erro! Marcador não definido.</b> |
| Figura 64-Testes utilizando a DWT com a wavelet mãe db8 .....                                     | 77                                  |
| Figura 65-Testes utilizando a DWT com a wavelet mãe haar .....                                    | 77                                  |
| Figura 66-Testes utilizando a EMD com 8 energias .....                                            | 78                                  |
| Figura 67-Testes utilizando a EMD com 2 energias .....                                            | 78                                  |
| Figura 68-Evolução da classe -1 ao longo dos 16 testes para cada método de pré-processamento..... | 79                                  |
| Figura 69-Evolução da classe -1 em cada eixo para cada método de pré-processamento.....           | 81                                  |



## LISTA DE TABELAS

|                                                                            |    |
|----------------------------------------------------------------------------|----|
| Tabela 1-Tipos de arquiteturas RNA .....                                   | 22 |
| Tabela 2-Topologias principais da RNA.....                                 | 23 |
| Tabela 3-Tipos de kernel mais utilizados .....                             | 24 |
| Tabela 4-Frequência de falha dos rolamentos.....                           | 47 |
| Tabela 5-Classificação da SVDD em cada eixo, utilizando a EMD.....         | 82 |
| Tabela 6-Classificação da SVDD em cada eixo, utilizando a DWT (db8) .....  | 82 |
| Tabela 7-Classificação da SVDD em cada eixo, utilizando a DWT (haar) ..... | 82 |
| Tabela 8-Experiências implementados .....                                  | 85 |
| Tabela 9-Resultados obtidos em cada experiência realizada .....            | 86 |



## LISTA DE ABREVIATURAS

|          |                                                                     |
|----------|---------------------------------------------------------------------|
| AG       | Algoritmos Genéticos                                                |
| AMR      | Análise Multi-Resolução                                             |
| BSS      | Blind Source Separation                                             |
| CART     | Classification And Regression Tree                                  |
| CHAID    | CHI-squared Automatic Interaction Detection                         |
| CWT      | Continuous Wavelet Tranform                                         |
| DWT      | Discrete Wavelet Transform                                          |
| ELM      | Extreme Learning Machine                                            |
| EMD      | Empirical Mode Decomposition                                        |
| Fast-ICA | Fast-Independent Component Analysis                                 |
| FCM      | Fuzzy C-Means                                                       |
| FFT      | Fast Fourier Transform                                              |
| FSVDD    | Fuzzy-Support Vector Data Description                               |
| HHT      | Hilbert-Huang Transform                                             |
| HMM      | Hidden Markov Model                                                 |
| HT       | Hilbert Transform                                                   |
| IMF      | Intrinsic Mode Function                                             |
| IML      | Interpolação Máxima Local                                           |
| KNN      | K-Nearest Neighbour                                                 |
| LDA      | Linear Discriminant Analysis                                        |
| OCSVM    | One Class Support Vector Machine                                    |
| PCA      | Principal component Analysis                                        |
| QUEST    | Quick, Unbiased, Efficient and Statistical Tree                     |
| RNA      | Redes Neurais Aritificiais                                          |
| RST      | Rough Set Theory                                                    |
| RSVDD    | Rough-Support Vector Data Description                               |
| SAFP     | Sliding Window Associated Frequency Pattern                         |
| SAFP-AD  | Sliding Window Associated Frequency Pattern-based Anomaly Detection |
| SIDL     | Shift Invariant Dictionary Learning                                 |

|        |                                            |
|--------|--------------------------------------------|
| SK     | Spectral Kurtosis                          |
| SNR    | Signal-to-Noise Ratio                      |
| SRM    | Structural Risk Minimization               |
| STFT   | Short Time Fourier Transform               |
| SV     | Support Vectors                            |
| SVDD   | Support Vector Data Description            |
| SVM    | Support Vector Machine                     |
| SVR    | Support Vector Regression                  |
| SW-ELM | Summation Wavelet-Extreme Learning Machine |
| WPT    | Wavelet Packet Transform                   |
| WT     | Wavelet Transform                          |

“We grow fearless by walking into fear”

Robin Sharma



# 1. INTRODUÇÃO

Atualmente o meio industrial encontra-se cada vez mais competitivo. Com a contínua e rápida evolução da tecnologia, para fazer face à concorrência e satisfazer um consumidor cada vez mais exigente, as indústrias tiveram que se adaptar e inovar os seus métodos de produção.

Surgiu assim a 4ª revolução industrial ou Indústria 4.0. Um modelo de indústria que utiliza, ao máximo, as tecnologias nos processos industriais, nomeadamente, as tecnologias de informação.

No início a única técnica de manutenção utilizada era a manutenção reativa ou por avaria, ou seja, a manutenção só era executada caso a máquina deixasse de funcionar. Com o objetivo de otimizar a produção, cortar custos e evitar grandes períodos de inatividade das máquinas, surgiu a manutenção preditiva. Neste tipo de manutenção a máquina é monitorizada e a manutenção ocorre antes da máquina ficar inativa.

A manutenção baseada na condição, que faz parte da manutenção preditiva, consiste em monitorizar o estado do equipamento durante a sua operação, considerando vários parâmetros como a vibração, temperatura, óleo de lubrificação, nível de ruído, entre outros.

Utilizando as tecnologias de informação é possível recolher e analisar dados, em tempo real, provenientes de vários sensores em máquinas, robôs, etc. Isto traz vários benefícios às empresas, como prognosticar ou diagnosticar falhas nos equipamentos, ajustar a produção, entre outros (Amaral, 2016).

Nascida em 1995, a Controlar é uma empresa dedicada ao desenvolvimento de hardware e software para a indústria, nomeadamente para a indústria automóvel.

A sua missão é desenvolver soluções de automação industrial eco inovadoras, com base em tecnologia adequada, garantindo soluções ambientalmente eficientes, com qualidade, segurança e inovação.

A presente dissertação faz parte de uma linha de investigação da empresa que tem como finalidade desenvolver uma solução para a monitorização de máquinas rotativas com capacidade de previsão da qualidade das peças produzidas. Para a realização dessa gestão inteligente, dados de vibração armazenados, extraídos de sensores aplicados numa máquina de produção de segmentos para motores de combustão interna, são analisados recorrendo a

técnicas estudadas. A empresa visa também utilizar as ferramentas desenvolvidas para a análise de dados de consumos energéticos de máquinas industriais.

Assim, um sistema de monitorização da condição capaz de diagnosticar falhas prematuras, é essencial numa empresa inserida na Indústria 4.0. É um sistema contínuo e autónomo que se traduz em grandes benefícios, pois permite a redução de custos de manutenção e evita perdas económicas e até humanas.

## **1.1 Motivação**

A manutenção preditiva, era inicialmente levada a cabo por analistas que periodicamente recorriam à sua experiência para detetar falhas nos equipamentos. Esta é uma boa abordagem, no entanto, para além de ter um custo associado, a empresa pode ter graves consequências quando o colaborador, responsável pela manutenção, não se encontra presente.

Hoje em dia, perante a Indústria 4.0, é possível adotar uma abordagem científica de manutenção com as vantagens das tecnologias de informação. Este método permite realizar uma monitorização mais eficiente e reduzir os custos de manutenção através da tomada de decisão com base na informação obtida através da supervisão dos elementos das máquinas industriais.

A gestão eficiente de uma máquina industrial pode passar pela monitorização contínua de diversos parâmetros essenciais ao bom funcionamento do equipamento. Estes são adquiridos através de sensores adequados para cada tipo de parâmetro.

O conjunto de dados resultantes da monitorização inclui qualquer possível parâmetro capaz de indicar o estado em que se encontra a máquina como, por exemplo, temperatura, vibração, som, corrente, entre outros.

A vibração é o parâmetro mais frequentemente utilizado pois, para além de ser facilmente adquirida, é muito sensível, possibilitando a deteção prematura de falhas antes do colapso da máquina e oferece um diagnóstico de falhas mais preciso, comparando com os outros parâmetros (P Konar, Sil, & Chattopadhyay, 2015).

Desta monitorização contínua resulta um grande volume de dados que é necessário interpretar. Hoje em dia, a indústria encontra-se muito interessada na exploração destes

grandes conjuntos de dados (Chen, Mao, & Liu, 2014), pois contêm informação valiosa, que leva à execução de tarefas ou tomada de decisões de uma forma mais fiável e eficiente.

Portanto, a abordagem científica (Ahmad & Kamaruddin, 2012) para previsão e monitorização autónoma, que recorre a técnicas de *data mining* e *machine learning* para a extração de conhecimento dos dados e para aprendizagem artificial do sistema é mais interessante, pois a sua utilização pode trazer grandes vantagens para a empresa, tanto a nível económico como ecológico.

Sendo assim, a implementação de um sistema de monitorização e previsão do estado operacional da máquina é extremamente útil para diminuir as falhas no equipamento, reduzindo os seus custos de manutenção e tempo de inatividade, levando à diminuição de perdas na produção e a uma melhor qualidade do produto (Ahmad & Kamaruddin, 2012).

## **1.2 Objetivos**

O processo de diagnóstico, essencial para efetuar a manutenção da condição, consiste em 3 tarefas principais:

### **1. Extrair dados que reflitam a condição de operação da máquina.**

Há vários parâmetros que podem ser obtidos, contudo, devido à sua fácil aquisição, disponibilidade e sensibilidade a falhas, a análise de vibração é a mais utilizada. Estes sinais são obtidos através de sensores colocados em várias posições estratégicas das máquinas.

### **2. Pré-processamento dos dados.**

Os sinais originais adquiridos não podem ser introduzidos diretamente nos métodos de diagnóstico pois são conjuntos de dados enormes, com redundância e ruído. Para serem eficientemente avaliados precisam de ser alvo de técnicas de processamento de sinal, de modo a extrair as suas características mais representativas. O tipo de técnica depende do tipo de sinal utilizado.

### **3. Diagnóstico de falha**

Nesta etapa são aplicadas técnicas capazes de transformar as características representativas do sinal em informação útil ou conhecimento.

Os objetivos desta dissertação foram estabelecidos com base nas tarefas essenciais na monitorização de uma máquina industrial. No entanto, como os dados a analisar já tinham sido adquiridos, a primeira tarefa não foi considerada.

Posto isto, os objetivos estabelecidos são os seguintes:

- Pré-processamento dos dados armazenados adquiridos da máquina industrial.
- Extração de conhecimento dos dados pré-processados, recorrendo a técnicas de *data mining*;
- Aplicação de técnicas de machine learning, para aprendizagem artificial do sistema;

### 1.3 Estrutura do documento

A presente dissertação está dividida em 7 capítulos principais que representam, adequadamente, a sequência de todo o trabalho realizado.

O capítulo 2 é dedicado a todos os fundamentos teóricos resultantes da revisão bibliográfica realizada, essencial para o desenvolvimento deste documento. Divide-se em dois subcapítulos principais, no primeiro são apresentados alguns métodos de processamento de sinal e no segundo são estudados métodos de extração de conhecimento e aprendizagem.

No capítulo 3 é descrita toda a metodologia desenvolvida, começando pela descrição dos dados e da ferramenta utilizada. Depois é feita uma descrição geral do desenvolvimento da dissertação, falando nas abordagens e nos métodos utilizados.

No capítulo 4 são apresentados os métodos implementados na primeira abordagem testada, a abordagem estatística.

O capítulo 5 expõem os métodos segundo uma abordagem de aprendizagem do sistema supervisionada.

No capítulo 6 encontram-se os dois últimos métodos desenvolvidos que seguem uma abordagem de aprendizagem do sistema não supervisionada.

No último capítulo, são apresentadas as conclusões obtidas, bem como sugestões para futuros trabalhos a desenvolver.

## 2. ESTADO DA ARTE

A monitorização da condição das máquinas industriais passa pela aquisição e análise de dados, em tempo real, com o objetivo de avaliar o estado operacional de cada máquina. Assim os problemas que causam as falhas nos equipamentos podem ser diagnosticados num estágio inicial dando a oportunidade de agir adequadamente antes que a máquina falhe (Baqqar, Ahmed, & Gu, 2011). Ao longo dos últimos anos, esta técnica tem ganho bastante interesse por parte de investigadores e da indústria.

Há várias formas de fazer a monitorização da condição dos equipamentos. Esta pode ser feita através da medição de vários parâmetros físicos, como a temperatura, o som, a corrente elétrica, entre outros (Ahmad & Kamaruddin, 2012; Baqqar et al., 2011). No entanto o parâmetro mais popular para este tipo de análise é o sinal de vibração, pois é facilmente obtido através de sensores, como, por exemplo, acelerómetros, e a sua análise permite o diagnóstico mais exato, quando comparado com outras técnicas (Pratyay Konar, Sil, & Chattopadhyay, 2015).

Estes sinais de vibração normalmente são adquiridos de motores (Pratyay Konar et al., 2015; Paulo, 2013; Yan, Member, Gao, & Member, 2006) caixas de engrenagens (Aherwar, 2013; Baqqar et al., 2011), ou rolamentos (Huang et al., 2016; Jia, Yuan, & Tang, 2011; Kamaras, Ili, & Dimitrakopoulos, 2016; Haining Liu, Liu, & Huang, 2011; Peng, Tse, & Chu, 2005; Plante, Stanley, Nejadpak, & Yang, 2016; Pourebrahimi & Mahmoodi, 2010; Tian, Morillo, Azarian, & Pecht, 2016; Zhao, Jin, Zhang, & Li, 2014; H. Zhou, Chen, Dong, & Wang, 2016), e são posteriormente analisados. Este tipo de análise inteligente consiste em três etapas principais: aquisição de dados, pré-processamento e a extração de conhecimento.

A aquisição de dados não é descrita neste documento, devido ao facto de os dados alvo de análise terem sido já obtidos pela empresa de acolhimento.

### 2.1 Pré-processamento

O pré-processamento é a etapa onde os dados vão ser limpos e transformados num formato adequado para a sua posterior exploração. Esta fase tem como objetivo a extração de características do sinal com o objetivo de identificar falhas. Cada falha tem uma frequência

característica ou “assinatura” correspondente que, normalmente, se encontra “escondida” no sinal de vibração.

Um bom pré-processamento é essencial para que as técnicas de *data mining e machine learning* tenham um bom desempenho, visto que o sinal de vibração contém muita informação e os dados recolhidos costumam ser redundantes e ter muito ruído.

Para preparar o sinal, normalmente utilizam-se técnicas de processamento de sinal que se dividem em 3 domínios: tempo, frequência e tempo-frequência.

No entanto, como estas técnicas possuem algumas limitações e problemas, os investigadores têm desenvolvido outros métodos de análise de dados como, por exemplo, a codificação *sparse*.

### 2.1.1 Domínio do tempo

Neste domínio, os algoritmos analisam diretamente as formas de onda e extraem características estatísticas, como o valor médio, desvio padrão, etc.

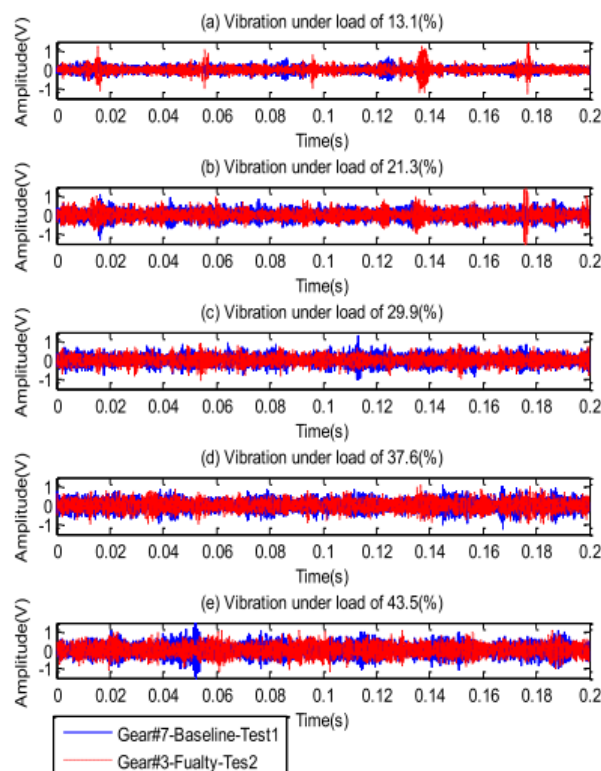


Figura 1-Sinais de Vibração de caixas de engrenagens em boas e más condições (Baqqar et al., 2011)

Na fig. 1 apresenta-se um exemplo de análise de sinais de vibração no domínio do tempo. Os sinais de vibração dizem respeito a 4 testes feitos a engrenagens em bom (azul) e mau (vermelho) estado com diferentes cargas. Embora se note alguma distorção, não dá instruções claras do tipo de falha (Baqqar et al., 2011).

Este tipo de domínio não é muito utilizado para análise de sinais de vibração. Não se adequa à análise de sinais não estacionários, porque não é capaz de fornecer informação suficiente à cerca das fontes de vibração (Javed, Gouriveau, & Zerhouni, 2015).

### 2.1.2 Domínio da frequência

O domínio da frequência é um dos métodos mais populares e eficientes para o diagnóstico de falhas, pois consegue identificar e isolar componentes de frequência. Embora não seja adequado para tratamento de sinais não estacionários (Javed et al., 2015), um dos métodos muito utilizados neste domínio, pela sua fácil implementação, é a *Fast Fourier Transform* (FFT). Alguns trabalhos utilizando este método são: Baqqar et al., 2011; Goyal & Pabla, 2015; Javed et al., 2015; Rashid, Amar, Gondal, & Kamruzzaman, 2016.

A análise do envelope, que utiliza a FFT, é também uma técnica muito utilizada para pré-processar dados de vibração (Kamaras et al., 2016).

#### *Fast Fourier Transform*

Esta técnica é utilizada para conversão de sinais de vibração no domínio do tempo para o domínio da frequência. O gráfico resultante da aplicação desta técnica apresenta os harmónicos do sinal (Aherwar, 2013) que, após a sua análise, permite a deteção de várias componentes de interesse, possibilitando o seu relacionamento com as componentes de falha (Rashid et al., 2016).

Na fig. 2 pode ver-se um exemplo ilustrativo do efeito da aplicação da FFT a um sinal acústico.

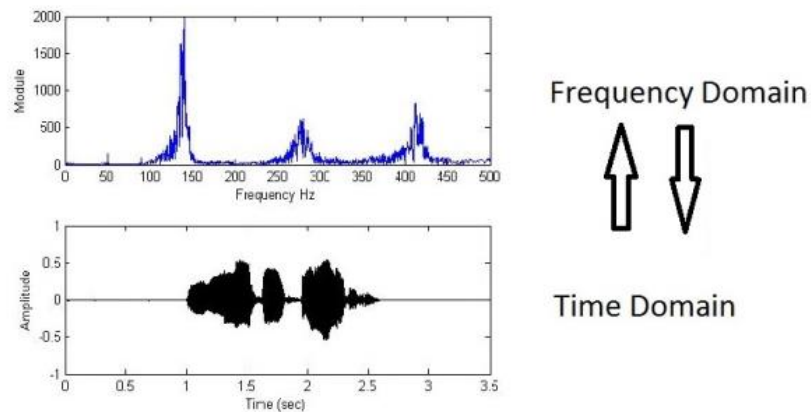


Figura 2-Aplicação da FFT (George Dallas, 2014)

Nesta técnica, o sinal é apenas considerado durante um intervalo de tempo, também designado de “*frame*” ou “*time window*”, como mostra a fig. 3. Assim é necessário escolher uma taxa de amostragem adequada para a digitalização do sinal e, também, para evitar componentes de frequência falsas resultantes do *aliasing* (Goyal & Pabla, 2015).

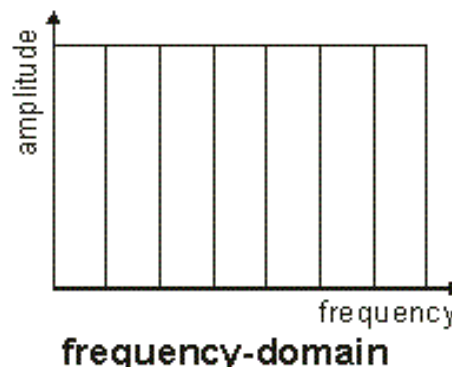


Figura 3-Frame ou Time Window utilizado no domínio da frequência (imagem adaptada de (Grosse & Reinhardt, 2002))

Para a escolha da taxa de amostragem, utiliza-se o teorema de *Nyquist* que enuncia que a frequência de amostragem deve ser pelo menos duas vezes a frequência mais alta que o sinal contém (Olshausen, 2000).

Como o sinal de vibração é dinâmico, este método apresenta algumas limitações, na medida em que não consegue retratar mudanças na frequência dos sinais ao longo do tempo (Goyal & Pabla, 2015), ou seja, embora se consiga detetar todas as frequências presentes no sinal, não se consegue saber o instante de tempo em que elas ocorrem (Valens, 1999), diz-se então que esta técnica tem apenas resolução na frequência.

A FFT é utilizada em (Rashid et al., 2016) para extrair as características mais importantes de dados de vibração de rolamentos.

### *Análise envelope*

A análise envelope, também conhecida como técnica de ressonância de alta frequência (Kamaras et al., 2016), é um método de referência no diagnóstico de falhas dos elementos rolantes das máquinas.

Esta técnica envolve a filtragem passa banda do sinal, a aplicação da FFT transformando o sinal para o domínio das frequências e o cálculo do envelope.

É uma técnica muito eficaz na identificação das falhas em estágios iniciais. Porém, a sua eficiência está dependente da escolha da banda com mais impulsos para a desmodulação. O método mais eficaz é determinar as características do filtro passa-banda, a frequência de corte e largura de banda, examinando as áreas do espectro de vibração com mudanças em dBs mais significativas do sinal original. Embora seja considerada a melhor abordagem, tem a desvantagem de necessitar de dados de vibração *à priori*, e que nem sempre estão disponíveis.

Com o intuito de resolver este problema, o *Spectral Kurtosis* (SK) tem sido utilizado para determinar a banda de frequência mais apropriada para a análise envelope (Kamaras et al., 2016).

#### **2.1.3 Domínio do tempo-frequência**

O domínio do tempo-frequência tem como objetivo analisar os sinais tanto no domínio do tempo como da frequência. São métodos muito poderosos na análise de sinais não estacionários, pois conseguem remover com sucesso o ruído e capturar componentes de frequência de interesse (Aherwar, 2013).

Neste domínio há muitos métodos, mas os mais utilizados são a *Short Time Fourier Transform* (STFT), *Continuous e Discrete Wavelet Transform* (CWT e DWT) (Javed et al., 2015; Pratyay Konar et al., 2015), *Wavelet Packet Transform* (WPT) e a *Hilbert-Huang Transform* (HHT), originada da *Hilbert Transform* (HT) (Pratyay Konar et al., 2015; Paulo, 2013).

O método *Blind Source Separation* (BSS) é muito utilizado para recuperar as diferentes fontes de sinal que constituem um sinal proveniente de um sensor. Foi muito utilizado em áreas como a biomedicina, telecomunicações, entre outros, mas só recentemente é que foi aplicado para monitorização e diagnóstico de máquinas industriais (Jing, Hai, & Kaixing, 2014).

### *Short Time Fourier Transform*

*Short Time Fourier Transform* (STFT) é o primeiro método de análise tempo-frequência (Safizadeh, Lakis, & Thomas, 2000), desenvolvido para proporcionar uma análise mais detalhada do que os métodos do domínio da frequência, até então utilizados.

Esta técnica permite dividir um sinal não estacionário em segmentos e depois aplica-lhes a FFT, como se pode ver na fig. 4.

Só é útil aplicar este procedimento em sinais onde esses períodos de observação repartidos (segmentos) sejam considerados estacionários, pois existem muitos sinais onde o conteúdo de frequência muda subitamente e, nestes casos, não se consegue estabelecer uma janela adequada para a sua análise (Selvakumar & Edwin Prakash, 2016).

A grande desvantagem da STFT é o facto de não conseguir oferecer uma boa resolução, simultaneamente, no domínio da frequência e do tempo. A resolução nos dois domínios está associada ao tamanho da janela que se escolhe para a sua implementação, ou seja, caso seja uma janela pequena tem-se uma melhor resolução no tempo, perdendo-se resolução no domínio da frequência, e vice-versa.

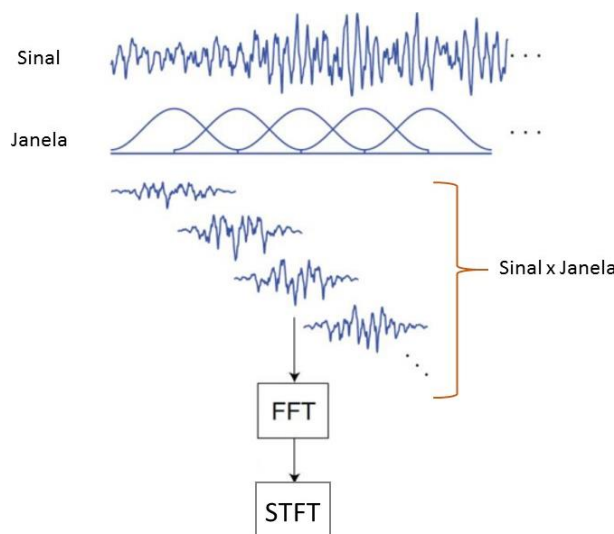


Figura 4-Método STFT (adaptado de (Gutierrez-Osuna, 2002))

Uma forma de contornar este problema é através da utilização das *wavelets*, tratadas mais à frente.

O comportamento da frequência do sinal ao longo do tempo, resultante da aplicação da STFT, pode ser observado através do espectrograma e a partir da análise deste.

No que diz respeito à utilização deste método na monitorização e diagnóstico de falhas dos componentes de máquinas industriais, como é um método de fácil implementação, rápido e eficiente, é muitas vezes escolhido (Safizadeh et al., 2000), mesmo havendo técnicas de processamento de sinal que fornecem uma análise mais rigorosa. Em (Hongmei Liu, Li, & Ma, 2016), a STFT é utilizada para pré-processar dados de som para extração das características das falhas, para depois serem introduzidos num classificador denominado de *deep learning network* e realizar o diagnóstico dos rolamentos.

### Wavelet Transform

Este método tem um grande potencial para detetar mudanças bruscas num sinal, é robusto e pouco sensível às mudanças nas condições de trabalho (e.g., quando se altera o local de trabalho da máquina). Consegue ainda eliminar ruído, comprimir dados e tem uma boa eficiência computacional (Goyal & Pabla, 2015).

Ao contrário da transformada de Fourier, que decompõe um sinal num conjunto de senos e cossenos, que são ondas infinitas, a transformada *wavelet* decompõe o sinal recorrendo a *wavelets*, caracterizadas como ondas pequenas e finitas que se desvanecem para 0, como se mostra na fig. 5 (George Dallas, 2014).

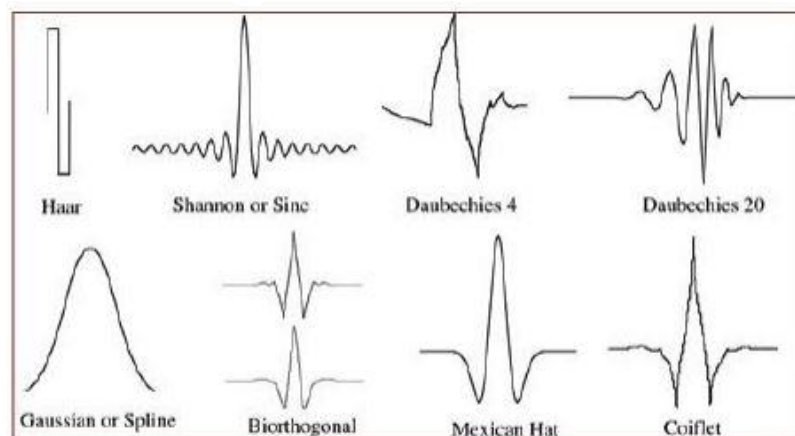


Figura 5-Tipos de Wavelets (George Dallas, 2014)

Para decompor o sinal, a transformada utiliza sempre o mesmo tipo de *wavelet*, também chamada de *wavelet* mãe, que funciona como uma janela escalada e modulada que desliza ao longo do sinal. Da decomposição resulta uma coleção de coeficientes do sinal, com diferentes escalas e translações (tempos) das *wavelets* filhas. Recorrendo a estes coeficientes e à *wavelet* mãe, é possível representar ou reconstruir o sinal (Haar, 1998). Devido a este conjunto de representações, esta análise chama-se também análise de multi-resolução (Valens, 1999).

Esta transformada permite representar um sinal do domínio do tempo, num sinal no domínio do tempo e da escala (frequência), como está na fig. 6, ou seja, altera uma função unidimensional para uma função bidimensional (Valens, 1999).



Figura 6-Representação tempo e escala resultante da aplicação da transformada de wavelets (Haar, 1998)

A escala é o comportamento das ondas em termos de frequência. Na fig. 7 está representado o comportamento da mesma *wavelet* em 3 frequências diferentes (Valens, 1999).

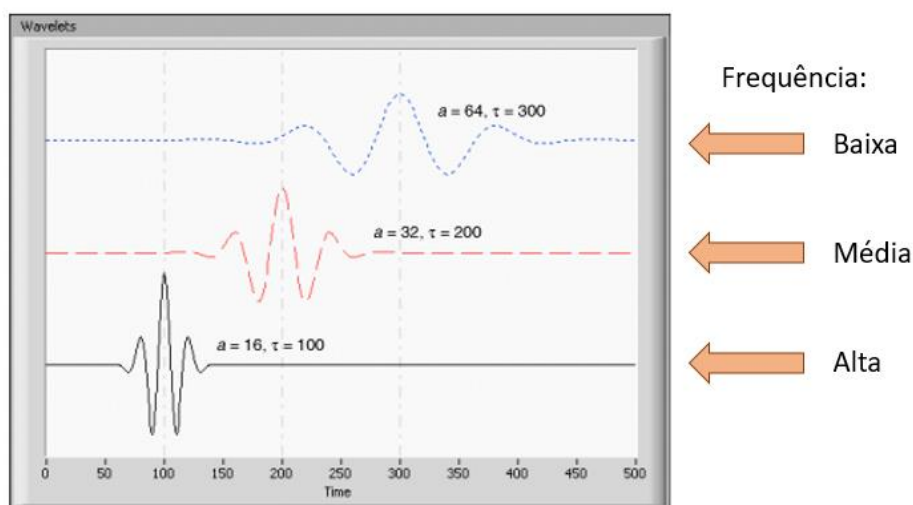


Figura 7-Wavelet com diferentes escalas (imagem adaptada de (George Dallas, 2014))

Como as *wavelets* são limitadas no tempo e na frequência, facilitam a análise do sinal e conferem uma maior resolução, tanto no domínio do tempo como da frequência.

No entanto, este método também tem alguns problemas como, por exemplo, interferência, distorção nas fronteiras e fugas de energia. Isto traz consequências para a interpretação dos resultados pois gera picos indesejados, podendo levar a análises incorretas.

Um dos problemas deste método que prevalece até hoje, é a escolha da *wavelet* mãe. A escolha é extremamente importante visto que a escolha errada ou inadequada da *wavelet* mãe pode reduzir a precisão no processo de detecção de falha (Loutas & Kostopoulos, 2012).

A técnica mais comum para a seleção desta *wavelet* base é escolhe-la através da sua semelhança em relação ao sinal em análise. No entanto este método foi provado como ineficaz por alguns artigos, (Li, Qu, & Liao, 2007; Loutas & Kostopoulos, 2012; Rafiee, Rafiee, & Tse, 2010), por essa razão, na maior parte dos casos a seleção é realizada por tentativa-erro.

A família *wavelet* mais popular é a Daubechies uma vez que tem um bom desempenho (Loutas & Kostopoulos, 2012) e é facilmente implementada, no entanto a escolha da *wavelet* mãe varia de aplicação para aplicação e, em (Loutas & Kostopoulos, 2012; Rafiee et al., 2010), foi provada a eficácia da *wavelet* Haar no diagnóstico de falhas em rolamentos e rodas dentadas.

### *Continuous Wavelet Transform*

A *Continuous Wavelet Transform* (CWT) fornece informação redundante e muito detalhada de um sinal no domínio do tempo e da frequência. São muito úteis para resolver problemas que envolvam a identificação de sinais e detecção de transientes (muito comuns nos sinais de vibração) (Fang & Luo, 2009), (Yan et al., 2006).

Este método consegue extrair informação detalhada do sinal mais eficientemente que a *Discrete Wavelet Transform*, contudo para isso, é necessário a escolha de uma escala e uma *wavelet* mãe adequadas (Pratyay Konar et al., 2015).

O parâmetro de escala é escolhido arbitrariamente ou por convolução (Goyal & Pabla, 2015). Se for selecionado aleatoriamente e não for adequado, pode-se perder informação. É uma escolha difícil, em que é necessário utilizar técnicas para a seleção da escala mais relevante que correlacione os coeficientes, escala e translação, obtidos após a aplicação da

transformada, com o sinal original (Pratyay Konar et al., 2015). O resultado da análise está intimamente dependente destas duas escolhas (Fang & Luo, 2009).

É um método adequado para a extração de características, as quais podem distinguir diferentes classes. No entanto, tem alguns inconvenientes. Além daqueles comuns às *wavelets*, não é o método ideal para a análise de grandes conjuntos de dados, visto que a computação de grandes amostras demora muito tempo, devido à sua redundância. Embora já tenha melhor resolução quando comparado com a transformada de Fourier, para altas frequências não tem uma boa resolução, simultaneamente, no domínio do tempo e da frequência (Pratyay Konar et al., 2015)(Peng et al., 2005).

Em (Pratyay Konar et al., 2015) é utilizada a CWT e a HT para fazer o processamento do sinal de vibração proveniente de motores em 7 condições diferentes. Após a extração das características de ambas as técnicas, foi realizado um estudo comparativo para selecionar a melhor técnica. A conclusão foi que ambas as técnicas, juntamente com métodos de *data mining*, tinham uma boa performance no diagnóstico de falhas do motor de indução.

### *Discrete Wavelet Transform*

A *Discrete Wavelet Transform* (DWT) é conseguida a partir da discretização da CWT. Esta técnica consiste na implementação da análise multi-resolução (AMR), que consiste no escalonamento e na *wavelet* associados a um filtro passa-baixo e um filtro passa-alto, respetivamente (Javed et al., 2015). Ou seja, a DWT é uma transformada *wavelet* associada a dois filtros iterativos, um passa-alto e um passa-baixo (Valens, 1999).

Na fig. 8 está representado como é efetuado o processo de decomposição da DWT. A banda de frequência, representada por  $x$ , com um tamanho de 16, vai passar pelos filtros passa-baixo e passa-alto um determinado número de vezes, dependendo do nível da transformada *wavelet* escolhido, que para este caso é 3.

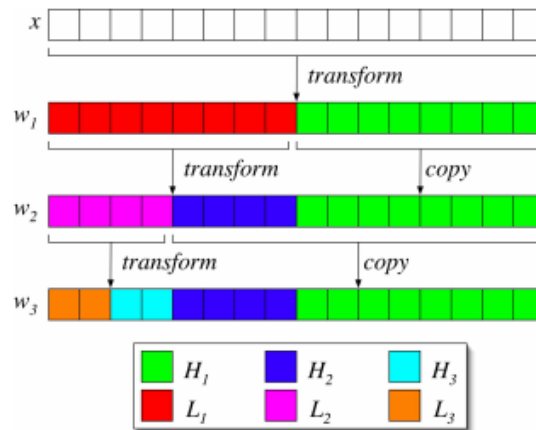


Figura 8-Resultado final do vetor  $x$ , depois de aplicar uma transformada discreta wavelet de nível 3 (Graps, 2004)

A cada nível,  $x$  é repartido em amostras, sendo cada amostra incrementada em dois a cada iteração. Primeiramente, passa pelo filtro passa-baixo até chegar ao fim do vetor. Os resultados adquiridos vão ser guardados nos primeiros 8 lugares. Seguidamente é realizado o mesmo procedimento, mas para o filtro passa-alto, e os resultados são guardados nos 8 lugares seguintes. Este processo pode ser observado na fig. 9 (Graps, 2004).

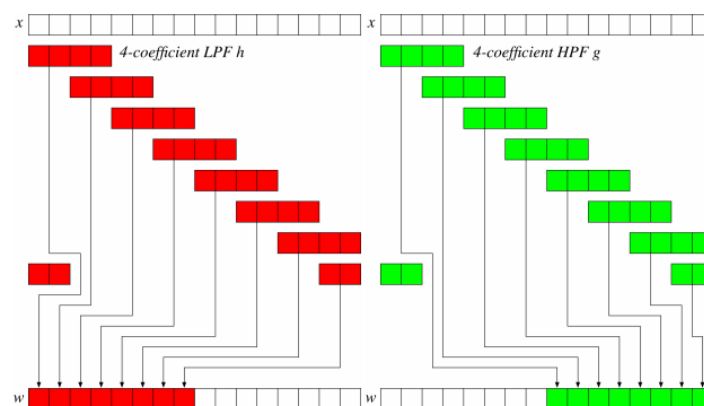


Figura 9-Procedimento realizado a cada nível da banda de frequência do método DWT (Graps, 2004)

Nesta transformada, o mais importante é construir bons filtros para que esta técnica seja aplicada com sucesso (Graps, 2004).

Em (Javed et al., 2015) a extração de características é feita em conjunto com funções trigonométricas, cuja vantagem é a transformação dos dados brutos em diferentes escalas, com o objetivo de obter características com melhores tendências e com pouca variabilidade comparando com as características clássicas. Para reduzir a variabilidade das características e atenuar o ruído é utilizado um método de suavização.

## Wavelet Packet Transform

Embora a DWT ofereça uma representação eficiente para muitos tipos de sinais, no processo de decomposição, como só é decomposta a parte de baixa frequência, perde-se informação que pode ser importante.

Um algoritmo capaz de superar esta desvantagem é *Wavelet Packet Transform* (WPT).

WTP é uma extensão da DWT que, implementando a AMR, permite a decomposição nível-a-nível do sinal (Dan & Wentao, 2014; Loutas & Kostopoulos, 2012).

O processo de decomposição é idêntico ao da DWT, no entanto, a cada iteração em vez de só ser decomposta a componente de baixa frequência (coeficientes de aproximação), também é decomposta a componente de alta frequência (coeficientes de detalhe), como podemos observar na fig. 10.

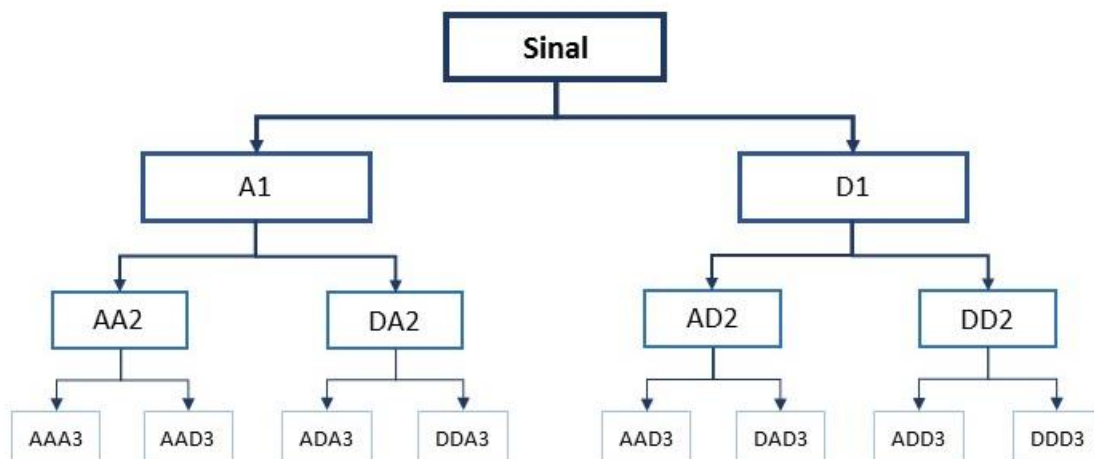


Figura 10-Decomposição em árvore da WTP de nível 3

No primeiro nível de decomposição o algoritmo vai decompor simultaneamente o sinal em dois pacotes, o coeficiente de aproximação A1 e o coeficiente de detalhe D1. No segundo nível cada pacote é dividido em dois sub-pacotes, de A1 resultam os coeficientes de aproximação e detalhe AA2 e DA2, de D1 advém os coeficientes AD2 e DD2. O processo continua até ao último nível de decomposição.

Como ambas as componentes são decompostas em todos os níveis, a análise *wavelet* por pacotes fornece uma resolução de frequência mais precisa do que a análise *wavelet* (Gokhale & Khanduja, 2010).

Assim, ao contrário do que acontecia na DWT que obteríamos  $(n+1)$  coeficientes, em que  $n$  é o nível de decomposição, na WPT obtém-se  $2^n$  nós.

A WPT tem sido utilizada no diagnóstico e prognóstico de máquinas (Dan & Wentao, 2014; Fei, 2017; Rajeswari, Sathiyabhama, Devendiran, & Manivannan, 2014; Xia, Xia, Wan, & Cai, 2012; Zhang, Wang, & Wang, 2013) e ainda de motores de indução (Khan, Radwan, & Rahman, 2007).

### *Hilbert Transform*

Esta técnica tem ganho muita popularidade em campos de análise de espectros de sinais, deteção de avarias, etc. (Fang & Luo, 2009). Resumidamente, esta transformada pode ser considerada como um filtro que desloca a fase do sinal em 90 graus. Um sinal analítico complexo é formulado, através do sinal original e da sua transformada de *Hilbert*, de onde pode ser extraída a amplitude em função do tempo (que é um índice da energia do sinal), e a fase, que está relacionada com a frequência de vibração (Fang & Luo, 2009).

Esta técnica também tem limitações, como o facto de esta transformada só ter significado físico quando é aplicada a sinais com uma única componente (Fang & Luo, 2009), o que não acontece na maior parte das aplicações em engenharia. Para resolver esta limitação, foi desenvolvida outra transformada, com base na transformada de *Hilbert*, muito popular, chamada de *Hilbert-Huang Transform* (HHT), na qual é proposto um método denominado *Empirical Mode Decomposition* (EMD), cujo fluxograma está na fig. 11.

Este algoritmo tem como objetivo a separação de componentes de um sinal, através de um método baseado em *splines* (interpolação cúbica) e extração de valores médios, num conjunto de mono-componentes, que são designadas por *Intrinsic Mode Functions* (IMF) (Fang & Luo, 2009; Paulo, 2013). Após a separação das mono-componentes, procede-se à análise do espectro de cada IMF (Goyal & Pabla, 2015).

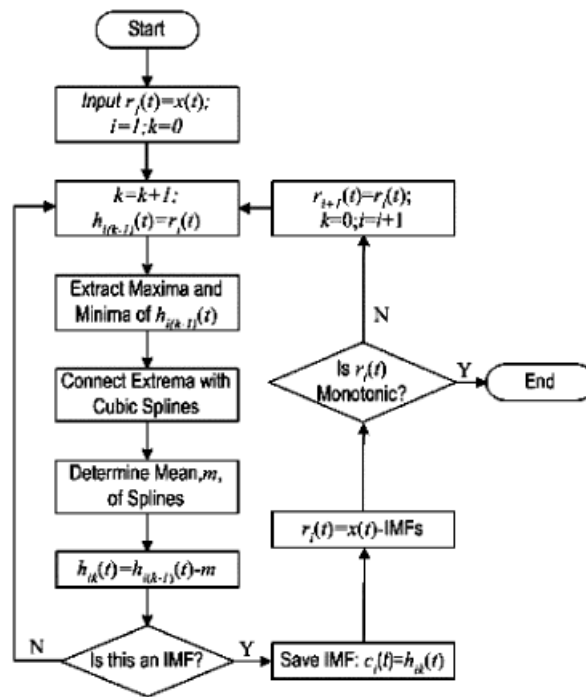


Figura 11-Fluxograma do método EMD (Yan et al., 2006)

A HHT tem sido muito utilizada em engenharia civil, na identificação de danos em estruturas. Em engenharia biomédica, tem sido utilizada para realizar estudos sobre arritmias cardíacas, a partir da análise dos batimentos cardíacos e da variação da frequência cardíaca, demonstrando a eficácia do método na detecção de transientes no sinal analisado (Yan et al., 2006).

Contudo, a HHT também tem desvantagens, todas relacionadas com o algoritmo EMD. Por essa razão os investigadores concentram-se em melhorar esse algoritmo (Fang & Luo, 2009). O EMD vai gerar algumas mono-componentes de baixa frequência, com baixa amplitude que são indesejadas. A primeira IMF gerada pode cobrir uma gama de frequências grande e, por isso, pode não conseguir satisfazer a condição de mono-componente. E, por fim, o método EMD não consegue separar algumas componentes do sinal com baixa energia, fazendo com que elas não apareçam no espectro (Fang & Luo, 2009; Peng et al., 2005).

No entanto, não necessita de qualquer convolução nem de outras operações que consomem muito tempo, conferindo-lhe uma rápida computação (Goyal & Pabla, 2015), (Peng et al., 2005).

Em (Fang & Luo, 2009) é utilizada uma versão melhorada da HT. Em teoria, a amplitude de Hilbert deve acompanhar os dados. Contudo devido ao efeito de amortecimento, a amplitude

aproximada de Hilbert é ligeiramente menor, como se pode ver na fig. 12. Como a HT envolve derivadas, pode ser muito sensível a essa diferença, mesmo que seja mínima. Para resolver este problema foi desenvolvida uma abordagem de interpolação máxima local (IML) para aproximar a amplitude de Hilbert ao seu extremo local.

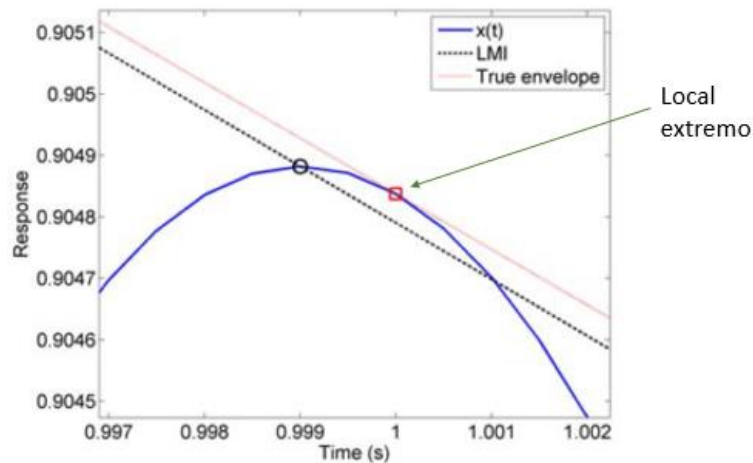


Figura 12-Envelope LMI (imagem adaptada de (Fang & Luo, 2009))

A HHT é utilizada em (Yan et al., 2006) para processar o sinal de um rolamento. Foi feita também uma comparação desta técnica com a Transformada de *Wavelet* e a STFT, concluindo que a HHT é a melhor técnica neste tipo de aplicações, devido às limitações apresentadas pelas outras técnicas.

Em (Peng et al., 2005) é feita uma comparação entre uma HHT melhorada e a *Wavelet Transform* (WT). A HHT foi melhorada através da utilização da *Wavelet Packet Transform* (WPT). Aos coeficientes gerados é aplicado o EMD, gerando as IMFs. Finalmente é selecionado um método de seleção de IMF para selecionar as funções úteis e descartar as restantes. No final a técnica HHT melhorada teve melhores resultados que a WT.

### *Blind Source Separation*

Os sinais de vibração reais normalmente têm uma taxa sinal-ruído, ou SNR (*Signal-to-Noise Ratio*), muito baixa, especialmente se a máquina se encontrar em ambientes muito ruidosos. Isto dificulta a captura de mudanças nos sinais de vibração caracterizadas por falhas ou degradação, nomeadamente no seu estado inicial, visto que, nessa fase essas mudanças são ainda mais subtis.

*Blind Source Separation* (BSS) é um método adequado para processar baixos SNR. Este considera que o sinal de vibração real é uma soma de sinais com fontes independentes. Desta forma, os sinais de vibração são entendidos como uma fonte, e o ruído é entendido como outra fonte independente da primeira. Assim, tendo em conta esta suposição, o algoritmo separa os sinais de fontes independentes, mesmo que os modelos destes sinais e número de fontes seja desconhecido (Huang et al., 2016). Uma representação gráfica deste processo está na fig. 13.

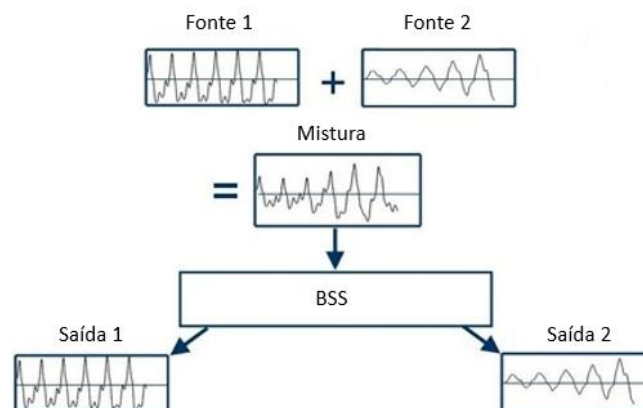


Figura 13-Método BSS (adaptado de (Han, Rafii, & Pardo, 2017))

Em (Huang et al., 2016), é utilizado um filtro BSS-SK (*Spectral Kurtosis*). Este filtro consiste primeiramente na separação do ruído dos sinais de vibração medidos, utilizando o Fast-ICA que é um dos métodos do algoritmo BSS. Após isso, é adotada a ferramenta, espectro curtose, para calcular os valores de curtose de cada espectro através da relação de impulsos e da frequência do sinal. Os valores de curtose são muito utilizados para descrever as características dos sinais de vibração.

Com este filtro foi possível obter melhores valores de curtose, o que levou a uma melhor detecção de falhas em estágios iniciais.

O espectro curtose foi utilizado várias vezes para analisar os sinais de vibração como, por exemplo, de rolamentos, utilizando também a análise envelope em (Kamaras et al., 2016), ou utilizando o classificador *K-Nearest Neighbour* (KNN) em (Tian et al., 2016).

## 2.2 Extração de conhecimento

A extração de conhecimento é a terceira parte do processo de diagnóstico inteligente. É nesta fase que se implementam as técnicas de *data mining* e *machine learning*, para extrair o conhecimento útil dos dados pré-processados. Este conhecimento é extraído através da análise de padrões escondidos nos dados, normalmente por algoritmos chamados de classificadores (Rashid et al., 2016).

Existe uma grande diversidade de famílias de algoritmos classificativos com as mais variadas particularidades. Neste documento vão ser apresentados os métodos estudados, agrupados pelo seu tipo de aprendizagem.

### 2.2.1 Algoritmos supervisionados

Os sistemas baseados em aprendizagem supervisionada constroem modelos onde as entradas são mapeadas para saídas desejadas (Hua, 2008). O objetivo é que o modelo construído permita antecipar uma saída para cada entrada. O sistema vai alterando o valor dos pesos, caso seja necessário, para que a sua saída seja o mais aproximada possível da saída modelo.

### *Redes Neurais Artificiais*

As Redes Neurais Artificiais (RNA) representam sistemas que recebem informação, armazenam-na, fazem o seu processamento, e transmitem o resultado final para o exterior.

É uma ferramenta excelente para explorar padrões, tendo a vantagem de deteção automática e identificação de condições de falha, sem necessitar de um conhecimento profundo do comportamento do sistema (Aherwar, 2013).

Esta técnica tem sido muito utilizada na classificação de padrões, modelação e tomada de decisão, tornando-a adequada para programas de monitorização (Aherwar, 2013).

Foram inspiradas no funcionamento do cérebro e, por isso, têm uma arquitetura idêntica à estrutura cerebral.

Numa rede deste tipo existem unidades de processamento, chamados de nós ou neurónios artificiais, ligados entre si e organizados em grupos (camadas). Estas ligações (sinapses) têm a

si associadas um peso, que vai ser multiplicado ao sinal transmitido, que representa a força do sinal enviado, se este for positivo, tratam-se de sinapses de excitação, caso contrário, tratam-se de sinapses de inibição.

Cada neurónio tem várias ligações de entrada/saída. Estes têm uma função de transferência/ativação aplicada à entrada para obter sinais de saída. O valor de entrada consiste na soma de todos os sinais de entrada multiplicados pelos respetivos pesos.

Para construir uma RNA tem que se ter em conta o processo de codificação da entrada e decodificação da saída, a definição da topologia da rede, o número e a ligação entre neurónios, a inicialização dos pesos da rede, e implementar um algoritmo adequado para o treino da rede (Santos & Azevedo, 2005).

A arquitetura da rede é definida pelo modo de ligação e organização dos neurónios. Existem várias arquiteturas, mas as mais utilizadas são as indicadas na Tabela 1:

*Tabela 1-Tipos de arquiteturas RNA*

| Tipo de rede      | Descrição                                                            |
|-------------------|----------------------------------------------------------------------|
| Totalmente ligada | Cada nó está ligado a todos os nós que constituem a rede.            |
| Camada única      | Redes que apresentam apenas duas camadas, a de entrada e a de saída. |
| Multicamada       | Redes constituídas por várias camadas paralelas entre si.            |

A primeira camada é designada por camada de entrada (*input*) e a última por camada de saída (*output*), as camadas intermédias são chamadas de camadas escondidas (*hidden layers*) e é aqui que ocorre o processamento. Pode observar-se esta arquitetura através da fig. 14.

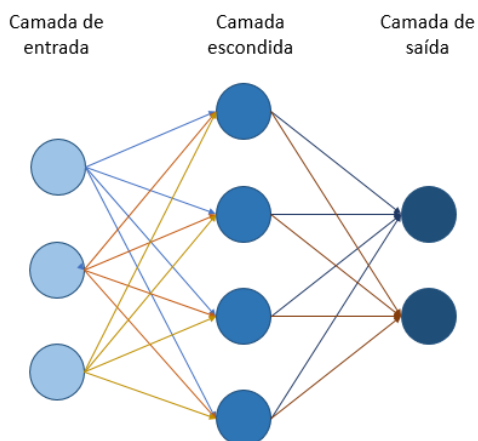


*Figura 14-Camadas de uma rede neuronal*

Há várias topologias de RNA, caracterizadas pelo modo de ligação dos nós, no entanto existem duas topologias principais, que estão descritas na Tabela 2:

*Tabela 2-Topologias principais da RNA*

| <b>Tipo de topologia</b> | <b>Descrição:</b>                                                                                                                                                                                                                                                 |
|--------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>FeedForward</i>       | Trata-se de uma rede acíclica, as ligações são unidirecionais e não há qualquer ligação entre neurónios das camadas anteriores e da mesma camada, apenas para as camadas seguintes. Esta topologia é mais utilizada, por ter cálculos simples e ser mais estável. |
| Recorrente               | Nesta rede, há ciclos de realimentação, ou seja, os neurónios de saída podem estar ligados aos de entrada. Aqui pode haver várias topologias.                                                                                                                     |

À semelhança do cérebro humano, as RNA também aprendem. Esta família de algoritmos, consoante as redes escolhidas, implementam três métodos de aprendizagem: aprendizagem supervisionada, aprendizagem não supervisionada e aprendizagem por reforço.

Em (Javed et al., 2015), foi utilizada uma técnica particular das RNA para realizar o prognóstico da máquina, a técnica *summation wavelet - extreme learning machine* (SW-ELM). Trata-se de um algoritmo que combina a RNA com a teoria *wavelet*, baseado em *extreme learning machine* (ELM). Este método diz-se eficaz na previsão/estimação das condições da máquina (Javed et al., 2015).

## Support Vector Machine

*Support Vector Machine* (SVM) faz parte de uma família de classificadores únicos e muito eficientes. Foi desenvolvida por Vapnik, para resolver problemas de classificação de padrões e de aproximação de funções não lineares (Chorowski, Wang, & Zurada, 2014).

Este método é baseado em princípios de minimização estrutural de risco (*structural risk minimization* - SRM), que supera alguns problemas das RNA (Dong, Yin, Tang, Chen, & Luo, 2014).

É uma técnica de classificação binária, pois só faz a classificação entre duas classes (classe +1 e -1). O seu paradigma principal é encontrar o melhor Hiperplano que separe as duas classes diferentes, de tal forma que a margem de separação entre classes seja máxima.

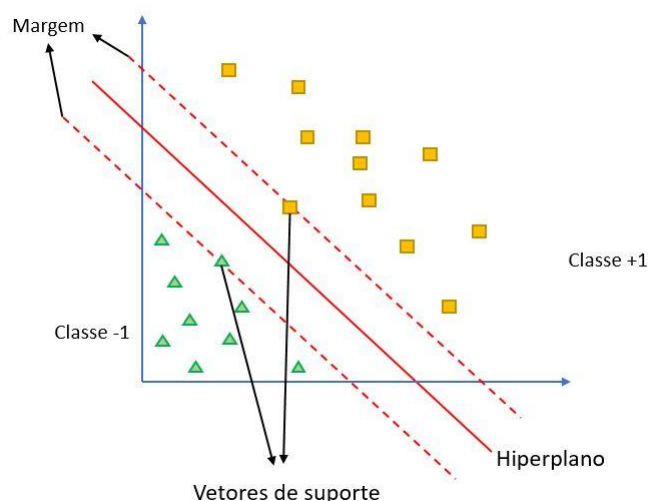


Figura 15-Classificador SVM

Na fig. 15 está retratada a forma como é realizada a classificação utilizando esta técnica. O limite entre as duas classes é descrito por um número limitado de dados, os vetores de suporte (*support vectors* - SV) (Joon, Eom, & Kim, 2005).

Em situações reais, encontrar dados linearmente separáveis é extremamente difícil, pois o sinal tem sempre uma componente de ruído associada, informação redundante, etc. Para processar dados não lineares, a SVM mapeia os dados de entrada num espaço de características de alta dimensão (*feature space*) utilizando uma transformação fixa não-linear, o kernel. Depois executa a classificação nesse espaço.

O kernel é um elemento chave na classificação pois a separação e o mapeamento das amostras dependem dele.

Tabela 3-Tipos de kernel mais utilizados

| Tipo de kernel | $K(x_i, x_j)$                   | Parâmetros     |
|----------------|---------------------------------|----------------|
| Polinomial     | $(\delta(x_i * x_j) + k)^d$     | $\delta, k, d$ |
| Gaussiano      | $\exp(-\sigma \ x_i - x_j\ ^2)$ | $\sigma$       |
| Sigmoide       | $\tanh(\delta(x_i * x_j) + k)$  | $\delta, k$    |

Na tabela 3 encontram-se os tipos de kernel mais utilizados e as suas respetivas funções. Os parâmetros apresentados são determinados pelo utilizador e têm que ser escolhidos adequadamente, pois os classificadores são muito sensíveis a eles.

A SVM é um algoritmo muito utilizado no diagnóstico e prognóstico de falhas em maquinaria industrial. Este método é utilizado sozinho ou em conjunto com outros métodos, com o objetivo de melhorar a sua performance.

Em (Dong et al., 2014) a SVM é utilizada, em conjunto com o *Markov Model*, para fazer o prognóstico do estado de rolamentos de uma máquina. O diagnóstico e monitorização destes componentes rolantes é feito em (Fangfang, Long, Xiucai, & Wenan, 2015; Hwang et al., 2015; Soualhi, Medjaher, & Zerhouni, 2015). No primeiro artigo é utilizado a HHT para fazer o pré-processamento dos dados, a SVM executa o diagnóstico do componente e a *Support Vector Regression* (SVR), um método baseado na SVM, é utilizado para estimar a sua vida útil. No segundo, o pré-processamento é realizado com a WPT. São retiradas as energias do sinal que vão constituir o vetor de características de entrada da SVM. No último artigo, a SVM é utilizada no diagnóstico de rolamentos para determinar a condição de motores de indução.

### *Fuzzy Logic*

Os Conjuntos Difusos (*Fuzzy Sets*) e a lógica difusa (*Fuzzy Logic*) têm um papel muito importante no processo de tomada de decisão em ambientes de incerteza, pois permitem a representação de raciocínio impreciso.

A lógica difusa assenta na premissa de que as experiências não podem ser classificadas como verdadeiras ou falsas por ser uma classificação incompleta. Assim, ao contrário dos sistemas lógicos clássicos, em que só se pode assumir dois valores (sim ou não), os atributos são exatos (p. ex. maior que, igual) e o modificador utilizado é a negação, na lógica difusa são utilizados conjuntos difusos (conjuntos que não tem limites definidos), ou seja, os valores podem tomar a forma de um conjunto ordenado (p. ex.: muito verdade, verdade, falso, muito falso), expressões booleanas ou um intervalo de valores, os atributos são subjetivos (p. ex.: bonito, feio) e existem vários modificadores (p. ex.: muito, mais, menos) (Santos & Azevedo, 2005).

Na fig. 16 está representado um exemplo comparativo da lógica difusa com a lógica clássica. A lógica difusa permite classificar as alturas dos indivíduos de uma forma mais precisa que a lógica clássica.

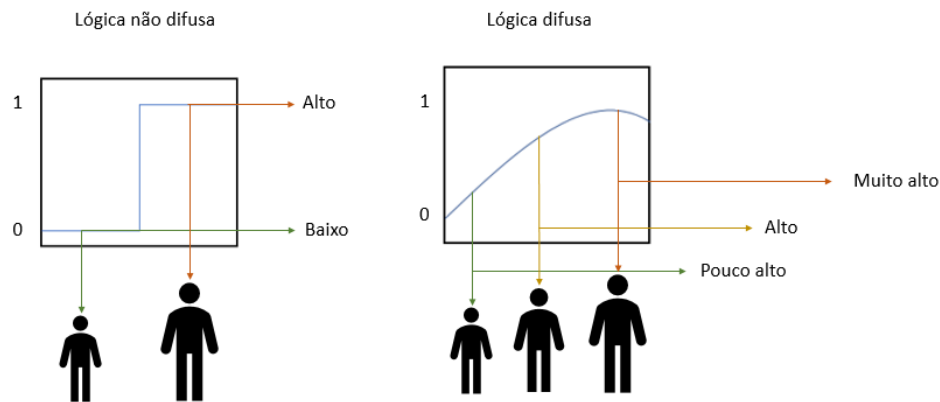


*Figura 16-Comparação do sistema de classificação clássico com a lógica difusa*

Esta é uma boa técnica para o tratamento de imperfeições no conhecimento, pois é capaz de detetar informação incompleta e converte-la para uma configuração numérica, facilmente manipulável pelos sistemas informáticos (Santos & Azevedo, 2005).

O seu uso na monitorização das condições das máquinas tem vantagens na medida em que estes algoritmos têm capacidade aproximada de raciocínio e conhecimento de linguagem embebida (Aherwar, 2013).

### *K-Nearest Neighbor*

Por vezes, no processo de classificação, a informação não se encontra presente, e os algoritmos utilizam a semelhança entre amostras para auxiliar o método de classificação, o *K-Nearest Neighbor* (KNN) tem sido utilizado nessas situações (Bagheri, Ahmadi, & Labbafi, 2010).

KNN é um algoritmo que armazena todas as classes disponíveis e classifica os novos casos baseando-se na similaridade entre eles e as classes reunidas (Dr. Saed Sayad, 2017). Este algoritmo faz parte dos métodos de indução de regras que é a técnica de *data mining* mais conhecida, cujo objetivo é a identificação de tendências, padrões ou dependências entre atributos ou valores num conjunto de dados através de probabilidades condicionais (Santos & Azevedo, 2005).

Quando é necessário tomar decisões sobre novos dados, este método calcula a distância entre estes dados e as amostras utilizadas no treino. Baseado num limite definido para a regra (k), as n amostras com as menores distâncias são seleccionadas e a classe com mais amostras é o resultado final. Este processo pode ser estudado a partir da fig. 17. Para  $k=3$ , dentro do raio do círculo encontram-se 2 triângulos e 1 quadrado, o resultado final vai ser a classe triângulo. Mas se  $k=5$ , como existem, no total 3 quadrados e 2 triângulos, o resultado final vai ser a classe quadrado.

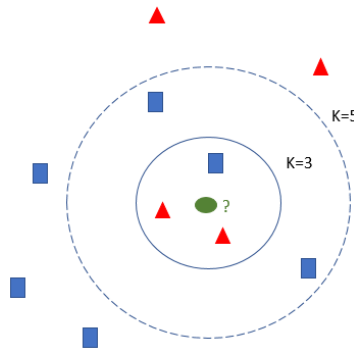


Figura 17-Exemplo do processo de decisão do algoritmo KNN

Esta técnica aparece, muitas vezes, associada a algoritmos de Árvores de Decisão, pois é muito utilizada na representação do conhecimento que está associado a esse algoritmo (Santos & Azevedo, 2005).

### Árvores de decisão

As Árvores de Decisão caracterizam-se pela representação de regras de classificação que seguem uma hierarquia de classes ou valores. Graficamente assemelham-se a uma árvore (Santos & Azevedo, 2005).

Na fig. 18 está representado o processo de uma árvore de decisão para o problema seguinte: a possibilidade de jogar ténis de acordo com o clima.

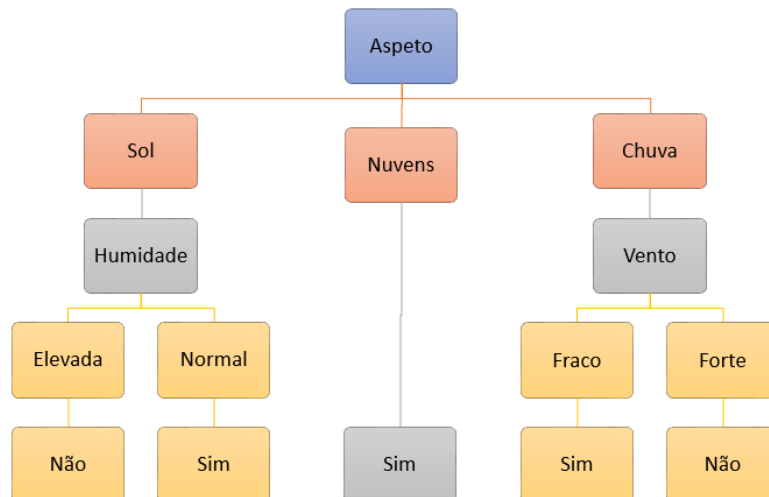


Figura 18- Árvore de decisão para o problema de jogar Tênis consoante o tempo

São um dos métodos mais populares de *data mining* e tem vários algoritmos como, por exemplo (Pourebrahimi & Mahmoodi, 2010): CART (*Classification And Regression Tree*), CHAID (*CHi-squared Automatic Interection Detection*), QUEST (*Quick, Unbiased, Efficient and Statistical Tree*) e C5.0.

Em (Pourebrahimi & Mahmoodi, 2010), é feita uma comparação entre esta técnica de *data mining* e as 3 anteriores, chegando à conclusão que o algoritmo C5.0 é o melhor destes modelos de classificação.

Outro algoritmo derivado das árvores de decisão foi proposto em (Rashid et al., 2016), o SAFP-tree. Este tipo de árvore que tem como base o algoritmo SAFP (*Sliding window Associated Frequency Pattern*) que captura o conhecimento importante do conteúdo recebido, em forma de pacotes, na atual janela de frequência.

Ainda neste artigo, o SAFP foi melhorado para ter a capacidade de detetar os diferentes tipos de falha, o que deu origem ao algoritmo SAFP-AD (*SAFP-based Anomaly Detection*). Estudos comprovam a eficiência e precisão desta técnica na descoberta de anomalias e falhas, utilizando os dados de vibração.

### *Fuzzy C-Means Clustering*

O *Fuzzy C-Means Clustering*, ou FCM, é um algoritmo inteligente de classificação baseado em lógica *fuzzy*. FCM é um dos métodos de *fuzzy clustering*, uma técnica que integra a análise

cluster com a lógica *fuzzy*. O objetivo do *fuzzy clustering* é dividir um conjunto de amostras em *clusters* com base na sua semelhança. Este método atribui, a cada amostra, graus de associação parcial, entre 0 e 1, que refletem a sua semelhança em relação aos *clusters* existentes, podendo uma amostra ter graus de associação para todos os *clusters*, não sendo “obrigada” a pertencer apenas a um único (Suganya & Shanthi, 2012). Este processo é iterativo e, por vezes, para se obter os resultados desejados é necessário alterar alguns parâmetros do algoritmo.

FCM dá um grau de associação a cada amostra com base na distância dessa amostra ao centro de um *cluster*. Quanto mais próxima a amostra se encontrar do centro do *cluster*, maior é o seu grau de associação a esse *cluster* específico. O somatório de todas as associações de uma amostra deve ser igual a 1 (Suganya & Shanthi, 2012).

Este algoritmo é muito utilizado em deteção de padrões, *data mining*, bioinformática, entre outros (Suganya & Shanthi, 2012). Quanto à avaliação da degradação de rolamentos, este algoritmo também é bastante utilizado. Os artigos (Cao, Ma, Zhang, Luo, & Yi, 2015; Lotfan, Salehpour, Adiban, & Mashroutechi, 2015; Yang, Guo, Zhang, & Liu, 2011) são alguns exemplos do envolvimento desta técnica no diagnóstico de elementos rolantes. Nos artigos o algoritmo é melhorado, ou associado a outras técnicas, conferindo bons resultados na deteção e classificação das falhas presentes nos rolamentos.

### 2.2.2 Aprendizagem não supervisionada

O método de aprendizagem não supervisionada permite ao sistema aprender através da deteção de padrões e de relações entre valores de entrada, auto adaptando-se sem conhecer as respostas corretas à priori. Este método de aprendizagem normalmente é utilizado quando não se dispõem de valores de saída, ou não há conhecimento de possíveis classificadores.

#### *One classe SVM e Support Vector Data Description*

Existem vários algoritmos, com diversas características, que tem como base a SVM. A *One Class SVM* (OCSVM) e a *Support Vector Data Description* (SVDD) são dois exemplos que fazem parte da família das SVMs cuja aprendizagem é não supervisionada.

## One Class Support Vector Machine

*One-Class Support Vector Machine* (OCSVM) é muito semelhante à SVM original. Mapeia os dados de entrada num espaço de características de grande dimensão e forma um hiperplano com o objetivo de distinguir a origem da classe positiva, chamada de classe *target*, de tal forma que a margem (distância do hiperplano à origem) seja máxima (Delft & Magnificus, 2001), como está retratado na fig. 19.

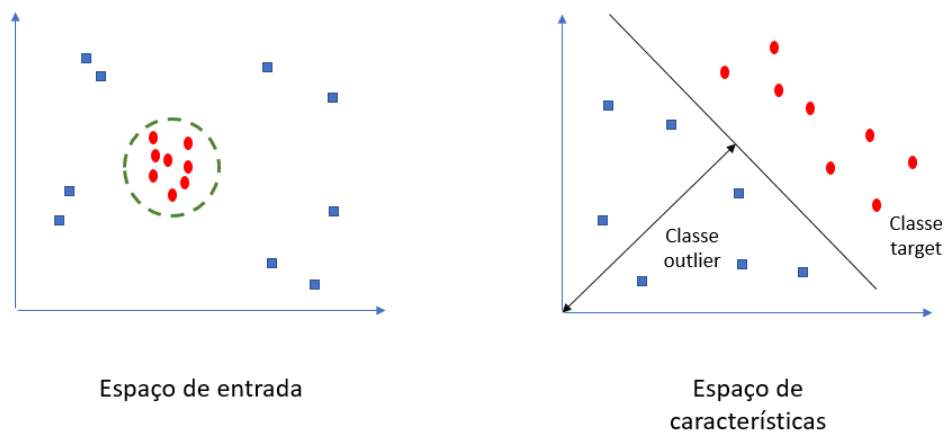


Figura 19-Método de classificação OCSVM

O resultado final é uma função binária que retorna +1 numa pequena região (região dos dados de treino) e -1 nos restantes locais, através da densidade de probabilidade das regiões (Roemer Vlasveld, 2013).

Quando é introduzido o conjunto de teste, todos os objetos que não são semelhantes aos da classe *target* (região +1) são considerados *outliers* e ficam fora do plano. Ou seja, o seu objetivo é distinguir exemplos normais, de eventos anormais através da observação dos dados de entrada.

A diferença entre a SVM original e a OCSVM é que a última apenas se dedica a uma classe, a classe *target*. Procura encontrar o padrão da classe *target* em vez de procurar semelhanças de duas classes.

Esta técnica mostrou bons resultados na deteção de falhas a partir de sinais de vibração, como pode ser visto em (Amir, Gul, & Khan, 2016; Deák, Kocsis, Vámosi, & Keviczki, 2014; Joon

et al., 2005; M., S., A., & L., 2012). Todos os artigos utilizam a OCSVM na classificação, mudando as técnicas de pré-processamento dos dados.

### *Support Vector Data Description*

*Support Vector Data Description* (SVDD) é idêntica à OCSVM, mas em vez de formar um hiperplano, forma uma hiperesfera, onde se encontram todos os elementos *target*, como se pode observar na fig. 20. O volume da esfera é minimizado com o objetivo de maximizar o número de *outliers* rejeitados pelo algoritmo.

SVDD oferece uma descrição mais flexível dos dados e tem ainda a vantagem de deixar incluir no conjunto de treino do algoritmo *outliers*, se disponíveis, com o objetivo de encontrar uma descrição dos dados mais eficiente (Delft & Magnificus, 2001). Estes devem ficar fora da esfera.

A diferença entre a SVDD e os outros métodos derivados da SVM é o facto de fornecer sempre um limite fechado à volta da classe *target*, tendo um foco especial nesta classe de dados. Como não estima a probabilidade de densidade da região completa, em contraste com o outro algoritmo estudado, consegue obter um limite melhor para conjuntos de dados limitados (Delft & Magnificus, 2001).

Como uma hiperesfera é um modelo muito rígido para mapear os dados, visto que estes não têm sempre a mesma representação no espaço, foram introduzidas as funções kernel, já descritas no subcapítulo das [SVM](#), para melhorar a classificação.

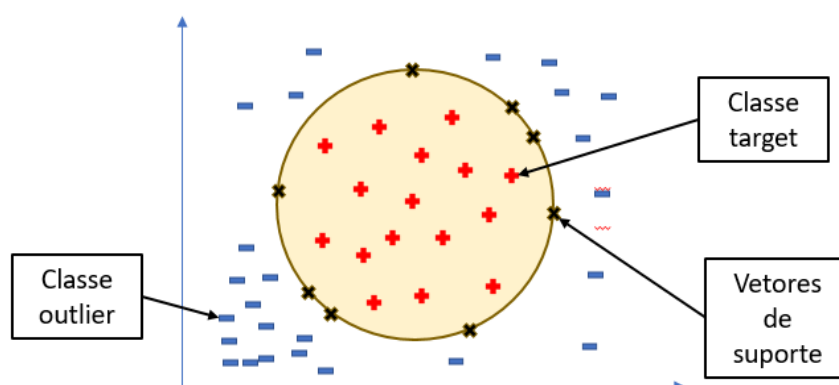


*Figura 20-Método de classificação da SVDD*

A capacidade de detecção de *outliers* da SVDD tem sido muito utilizada em várias aplicações como detecção de falha em bombas, reconhecimento facial e de voz, imagem médica (Benkedjouh, Medjaher, Zerhouni, & Rechak, 2012) e na avaliação e previsão da degradação de rolamentos (Benkedjouh et al., 2012; Pan, Chen, & Dong, 2009; J. Zhou, Guo, Zhang, Xu, & Li, 2016). Em (Shen, He, Chen, Sun, & Liu, 2012; Zhu, Zhang, & Zhu, 2013) são desenvolvidos dois métodos híbridos que provêm da SVDD, a *Fuzzy SVDD* (FSVDD) à qual são associados os graus de associação da lógica *fuzzy*, e a *Rough SVDD* (RSVDD) que agrega os princípios da teoria *rough set* com a SVDD para superar os problemas das funções de associação da FSVDD. Em ambos os artigos é realizado o diagnóstico de degradação dos rolamentos.

### 3. METODOLOGIA

Neste capítulo vai ser apresentado e descrito todo o processo de desenvolvimento, sendo inicialmente abordados os tipos de dados analisados. Posteriormente é feita uma elucidação da ferramenta utilizada. Depois são apresentadas, de uma forma geral, as técnicas estudadas e desenvolvidas, divididas em 3 fases: filtragem, pré-processamento, extração de características, e classificação e aprendizagem.

Esta dissertação foi desenvolvida iterativamente. Foram implementadas várias combinações de algoritmos, para alcançar uma combinação ótima que vá de encontro aos objetivos propostos.

Com o progresso do desenvolvimento, estas combinações vão se tornando mais complexas devido à natureza de cada algoritmo envolvido.

#### 3.1 Dados

Para que a monitorização de uma máquina seja executada com sucesso, os sensores utilizados e a sequência de aquisição de dados são fatores críticos. A precisão do sensor, a sua utilização adequada e sitio de montagem determinam a qualidade e validade dos dados adquiridos (Mobley, 2001).

Assim, no desenvolvimento deste projeto foram utilizados três conjuntos de dados diferentes, descritos nas subsecções seguintes.

##### 3.1.1 Dados de referência

Os dados de referência, extraídos do site do Centro de Sistemas Inteligentes de Manutenção (*Center on Intelligent Maintenance System*) da Universidade de Cincinnati, foram utilizados com o objetivo de assegurar que os algoritmos implementados estavam a produzir uma boa classificação.

O pacote de amostras é constituído por 3 conjunto de dados, onde cada um representa uma experiência teste-falha, executada numa plataforma especialmente projetada para o efeito, em condições de carga normal (Qiu, Lee, Lin, & Yu, 2006). Esta plataforma e o seu respetivo esquema encontram-se na fig. 21.

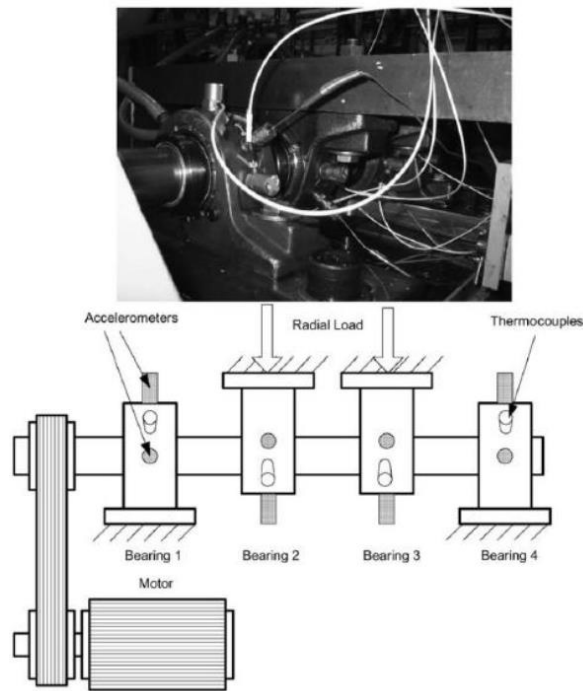


*Figura 21-Plataforma de teste e esquema de posicionamento dos acelerômetros (Qiu et al., 2006).*

Cada conjunto de dados é caracterizado por arquivos que individualmente contêm sinais de vibração com duração de 1 segundo gravados em períodos de tempo específicos. Os dados foram adquiridos recorrendo a um programa em LabView (Qiu et al., 2006).

Este conjunto de dados foi utilizado num artigo (Huang et al., 2016), que foi implementado durante o presente trabalho, em que ocorre uma falha na calha interior no final da experiência do teste-falha. Apenas foi utilizado o eixo dos X que consiste em 2156 amostras. Neste total de amostras não se sabe ao certo o momento em que ocorre a falha, apenas que será nas últimas amostras, pelo que a sua utilização na sequência temporal é imperativa.

### 3.1.2 Dados disponibilizados pela empresa de acolhimento

Para avaliar a performance de todos os algoritmos e técnicas implementadas na monitorização de máquinas industriais, a empresa Controlar disponibilizou 268 amostras de dados de vibração de uma máquina inserida num ambiente industrial com elevado ruído, adquiridos em tempo real do seu funcionamento por um acelerómetro tri-axial (X, Y e Z), que se encontra na fig. 22.

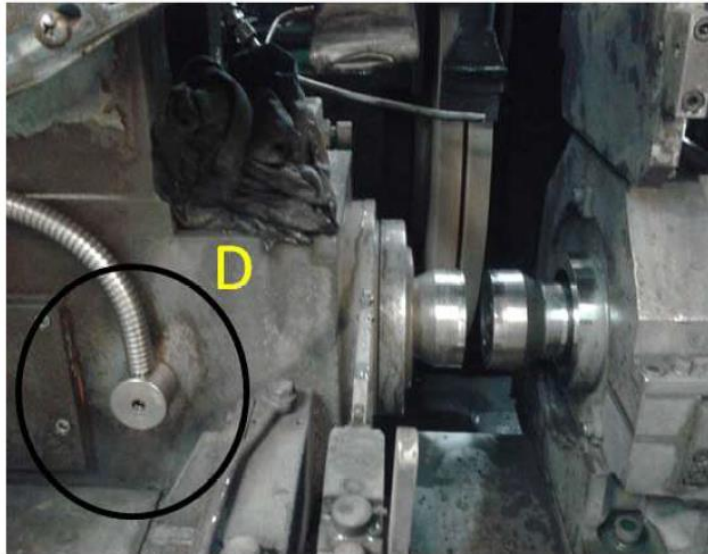


*Figura 22-Posicionamento do sensor (D) na máquina teste*

Desses dados sabe-se apenas que 9 amostras contêm falhas manipuladas, 67 amostras representam o bom funcionamento da máquina e para as restantes, ou seja, 192 amostras não se conhece o estado da máquina.

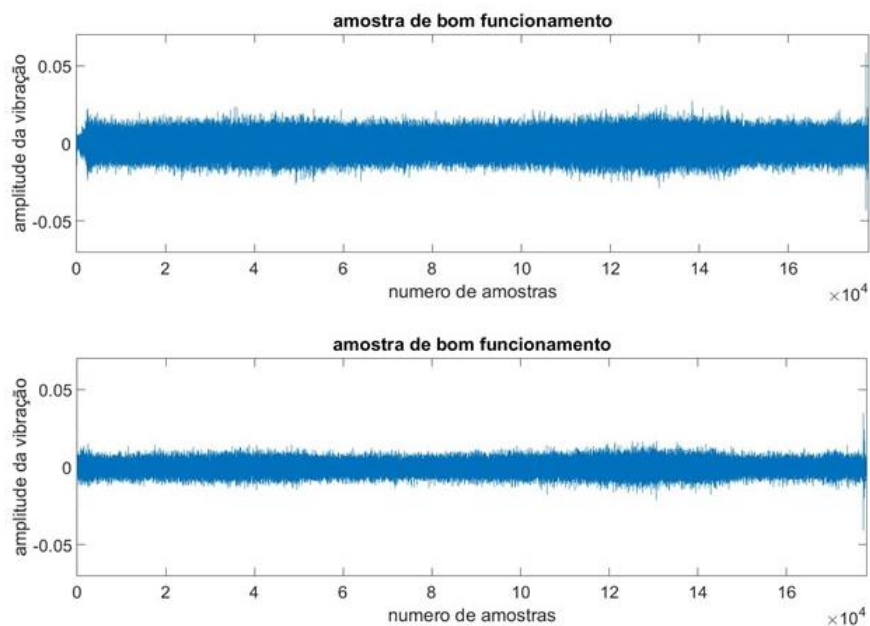


*Figura 23-Exemplos de amostras de vibração que caracterizam o mau funcionamento do equipamento*

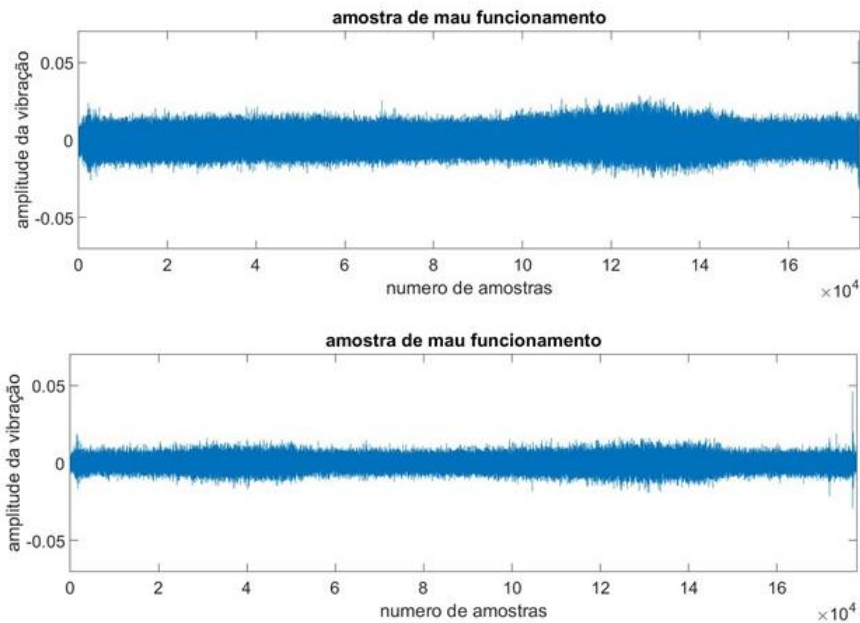


*Figura 24-Exemplos de amostras de vibração que caracterizam o bom funcionamento do equipamento*

Nas fig. 23 e 24, estão representadas 4 amostras. Duas fazem parte do conjunto de 9 amostras que contêm falhas manipuladas, e as outras duas amostras representam o bom funcionamento da máquina.

Observando estas duas figuras, não se consegue retirar informação útil a cerca de ocorrências de falhas no equipamento ou a sua natureza. Por esta razão, para estudar estas amostras de vibração são necessários métodos de análise mais poderosos, pois através da análise visual do sinal não é possível extrair conhecimento relevante sobre o problema alvo.

### 3.1.3 Dados que caracterizam falhas específicas

Foram retirados da página do Centro de Dados de Rolamentos da universidade de *Case Western Reserve*, o qual dá acesso a conjuntos de dados de vibração que caracterizam falhas específicas de rolamentos, e o seu funcionamento normal.

Os dados de vibração foram adquiridos através de acelerómetros colocados nos rolamentos de um motor elétrico.

Na fig. 23 encontra-se o motor utilizado para a recolha dos dados.

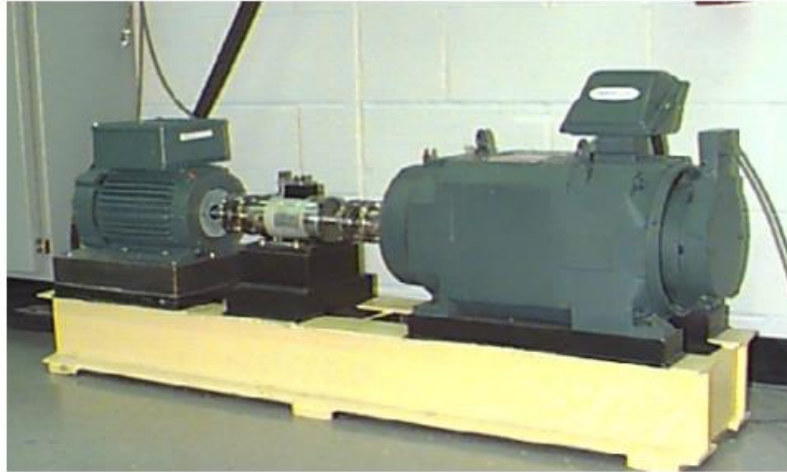


*Figura 25-Motor elétrico utilizado na aquisição de dados de vibração (University, 2017)*

A direção e posicionamento dos sensores no objeto alvo de medição é muito importante para se obter dados informativos valiosos, no que diz respeito à monitorização da máquina.

Como a resposta dos objetos mecânicos a vibrações é complexa, pode-se esperar níveis de vibração significativamente diferentes mesmo que as medições sejam feitas em dois pontos de medição adjacentes no mesmo elemento da máquina (Units & Influences, 1982).

## **3.2 Ferramentas utilizadas**

Antes de iniciar a fase de desenvolvimento, procedeu-se à escolha da ferramenta de software para implementar todos os métodos desenvolvidos nesta dissertação. A ferramenta escolhida foi o Matlab da MathWorks.

### **3.2.1 Matlab**

O Matlab é uma ferramenta de software versátil, poderosa e com ambiente interativo que possibilita a computação numérica, visualização e manipulação de dados. A sua linguagem de programação é de alto nível o que facilita a sua utilização.

Internamente, o software fornece um conjunto funções de carácter básico, mas para além destas, existem bibliotecas com funções adicionais (*toolboxes*) que proporcionam soluções para problemas simples e complexos. Com estas bibliotecas é possível realizar o processamento de sinais e a análise de dados.

Esta ferramenta foi escolhida devido à sua sintaxe simples, intuitiva e ao seu imenso repositório de *toolboxes* disponíveis de técnicas e algoritmos de processamento de sinal, *Data Mining* e *Machine Learning*.

Uma vantagem adicional da utilização do Matlab é a possibilidade de integrar o código na ferramenta de software LabView, caso o sistema de recolha de dados tenha sido desenvolvido nesta plataforma.

### 3.3 Abordagem e Métodos

Os conceitos de *data mining* e *machine learning* são cada vez mais conhecidos e existem muitas técnicas diferentes para executar estes conceitos.

Com o intuito de estudar todo o processo de monitorização e as técnicas mais adequadas ao objetivo proposto, o desenvolvimento deste trabalho foi realizado iterativamente, através da implementação e teste de vários métodos de pré-processamento e técnicas de classificação, tendo como base toda a literatura estudada. Quando um método implementado apresentava limitações e não se adequava ao problema, eram estudados, escolhidos e implementados novos algoritmos até ser obtida uma solução que satisfizesse o objetivo principal.

No desenvolvimento foram utilizados dois conjuntos de dados, os dados de referência (3.1.1) e os dados cedidos pela empresa de acolhimento (3.1.2).

Inicialmente, a performance dos algoritmos era testada recorrendo aos dados de referência. Se os testes de performance demonstravam bons resultados, eram introduzidos os dados da empresa Controlar, para verificara sua eficácia em dados de vibração de uma máquina industrial real.

Foi ainda utilizado um conjunto de dados com falhas específicas apenas num teste realizado.

As técnicas estudadas foram divididas em 3 conjuntos, considerados fundamentais no desenvolvimento de um sistema de deteção de falha. Estes são brevemente descritos de seguida.

### 3.3.1 Filtragem

Na aquisição de dados de vibração de uma máquina industrial, o sinal obtido é composto por várias fontes (de vibração), como, por exemplo, dos rolamentos, motor, etc. Para além destas fontes, a máquina também está sujeita a vibrações de outras máquinas vizinhas e ruídos do ambiente em que se encontram inseridas.

Assim, os sinais adquiridos contêm muito ruído dificultando a sua análise, principalmente na deteção de leves mudanças no padrão de vibração, pois encontram-se camufladas pelo ruído.

Assim é importante retirar o ruído do sinal para que a sua análise seja eficiente. Para isso, neste trabalho foram utilizadas duas técnicas diferentes: Blind Source Separation (BSS) e Wavelet Denoise (WD).

### 3.3.2 Pré-processamento e Extração de características

Atenuando o ruído do sinal de vibração, este encontra-se pronto para ser pré-processado. O pré-processamento é muito importante pois permite retirar características específicas do sinal, que refletem a evolução das falhas, melhorando a sua análise.

As técnicas utilizadas para pré-processar dos dados de vibração foram as seguintes: curtograma rápido associado a uma análise envelope, *Discrete Wavelet Transform* (DWT), *Wavelet Packet Transform* (WPT) e *Empirical Mode Decomposition* (EMD).

Através destas técnicas foram extraídas as seguintes características estatísticas:

- Valor eficaz (RMS)

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \quad (1)$$

- Média

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n |x_i| \quad (2)$$

- Variância

$$Var = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \quad (3)$$

- Desvio padrão

$$DP = \sqrt{\frac{1}{n-1} \sum_{i=1}^n |x_i - \bar{X}|^2} \quad (4)$$

- Valor pico-a-pico

$$V_{pp} = \left( \frac{1}{n} \sum_{i=1}^n X_{pi} \right) * 2 \quad (5)$$

- Fator de pico (*Crest Factor*)

$$CF = \frac{\left( \frac{1}{n} \sum_{i=1}^n X_{pi} \right)}{RMS} \quad (6)$$

- Curtose (Kurtosis)

$$Kur = \frac{1}{n} \frac{\sum_{i=1}^n x_i^4}{RMS^4} \quad (7)$$

- Assimetria (Skewness)

$$Skew = E \left[ \left( \frac{x - \mu}{\sigma} \right)^3 \right] \quad (8)$$

Por vezes, através de comandos do Matlab, extraíram-se também as energias das amostras de vibração para estudo.

Antes de introduzir as características representativas dos sinais nos classificadores, é essencial normalizá-las para que sejam tratadas de igual forma pelo classificador.

Foram implementados dois tipos de normalização. O primeiro procedimento é a normalização padrão onde as características são quantificadas entre 0 e 1, através da média e do desvio padrão.

$$X_{norm} = \frac{(X - \bar{X})}{DP(X)} \quad (9)$$

O segundo procedimento utilizado foi retirado de (M. et al., 2012), que consiste em 3 passos:

1. Calcula-se a energia total para n decomposições (ou IMFs).

$$E_n = \int_{-\infty}^{+\infty} |c_n(t)|^2 dt \quad (10)$$

2. Cria-se um vetor de características com as energias extraídas anteriormente.

$$F = [E_1, \dots, E_n] \quad (11)$$

3. Normaliza-se o vetor de características dividindo F pela energia total de todas as decomposições E.

$$E = \left( \sum_{i=1}^m |E_n|^2 \right)^{\frac{1}{2}} \quad (12)$$

Sendo o novo vetor de características:

$$F' = \left[ \frac{E_1}{E}, \dots, \frac{E_n}{E} \right] \quad (13)$$

### 3.3.3 Classificação e Aprendizagem do Sistema

Após o tratamento adequado dos dados, é necessário analisá-los para ser possível extrair informação dos mesmos e tomar decisões com base nessa informação.

Como ferramenta de análise e aprendizagem são utilizados classificadores. Estes são algoritmos capazes de descobrir padrões e reconhecer as falhas já “aprendidas” a partir das características extraídas.

Os classificadores utilizados foram os seguintes: Support Vector Machine (SVM), Fuzzy C-Means Clustering (FCM), One-Class SVM e Support Vector Data Description.



## 4. ABORDAGEM ESTATÍSTICA

Os próximos 3 capítulos vão conter uma descrição mais detalhada do desenvolvimento de cada implementação, e a demonstração e discussão de resultados. Estes são apresentados de acordo com a evolução dos métodos implementados e do estudo realizado.

No processo de desenvolvimento iterativo explicado no começo do capítulo 3, inicialmente optou-se por uma abordagem estatística. Foram executados dois testes com algoritmos de processamento de sinal diferentes apresentados de seguida.

A normalização utilizada nestes dois testes foi a normalização standard.

### 4.1 *Blind Source Separation* + Curtograma Rápido e análise envelope

No começo da etapa de desenvolvimento implementou-se um método descrito em (Huang et al., 2016), que sugere uma técnica nova de pré-processamento de sinal.

O artigo utiliza a técnica BSS para fazer a filtragem e a combinação do método curtograma rápido com a análise envelope para realizar o pré-processamento do sinal. O diagrama desta implementação está na fig. 26.

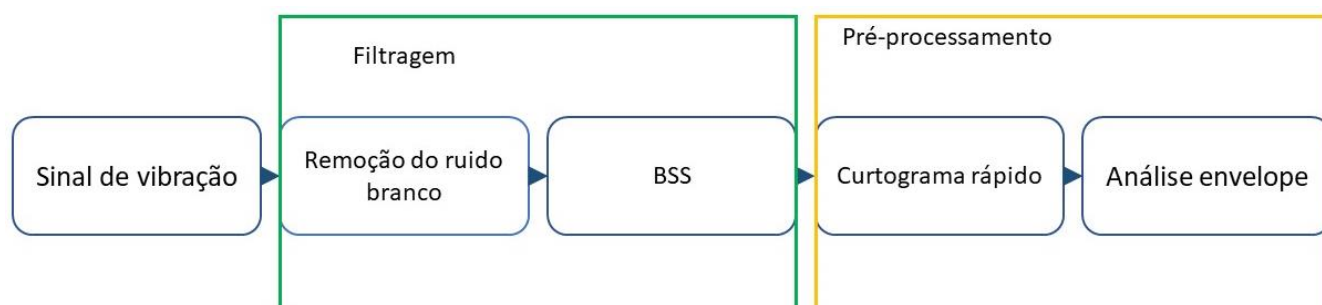


Figura 26-Diagrama da implementação BSS + Curtograma rápido e análise envelope

#### 4.1.1 *Blind Source Separation*

Na aquisição de dados de vibração de uma máquina industrial, o sinal obtido é composto por várias fontes de vibração, como o ambiente, rolamentos, motor, etc. O método *Blind Source Separation* (BSS) realiza a separação do sinal de vibração em várias fontes, sem saber *à priori* quantas fontes existem no sinal e a sua estrutura misturada.

Uma das particularidades deste algoritmo é que, ao efetuar a filtragem, este altera a fase e amplitude do sinal original, não sendo adequado para a extração de algumas características

clássicas, como a RMS, média, variância, etc. No entanto, características como a curtose, que é a quantidade de picos existentes no sinal (Wang & Markert, 2015), não é afetada pelas alterações na fase e na amplitude do sinal (Huang et al., 2016).

Antes de realizar o processamento para extrair a característica da curtose é boa prática remover do sinal o ruído branco. Este procedimento vai maximizar os valores de curtose e elimina o efeito de algumas componentes estacionárias (Sawalhi & Randall, 2005).

Posto isto, o BSS é utilizado para separar o sinal de vibração do ruído. O algoritmo utilizado para implementar o BSS foi o *Fast-ICA* (*Independent Component Analysis*), como pode ser visto na fig. 28. Ao sinal normalizado n1 é aplicado o algoritmo *Fast-ICA*, ficando o resultado final na variável “Sig\_BSS”. Os resultados encontram-se na fig. 29.

```
[Sig_BSS] = fastica(n1);
```

Figura 28-Função Matlab utilizada na implementação do Fast-ICA

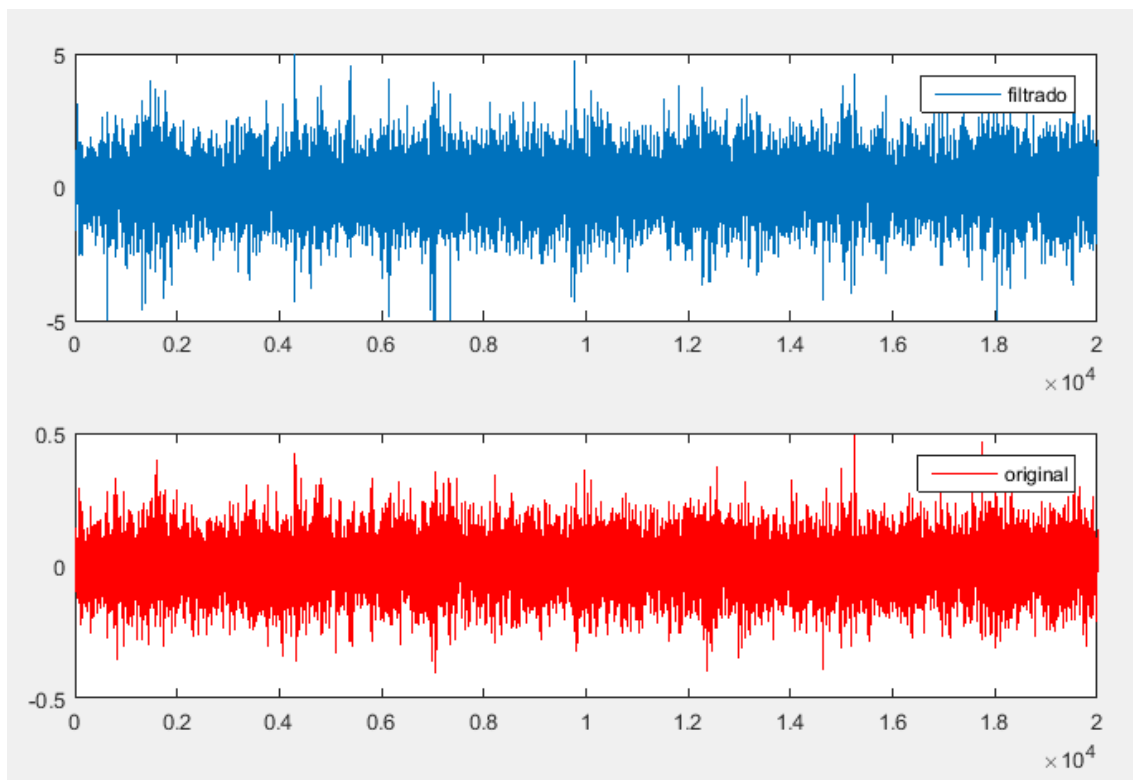


Figura 29-Sinal de vibração filtrado, utilizando o algoritmo Fast-ICA

Na fig. 29 estão representados dois sinais de vibração: o sinal a vermelho é o sinal original e a azul o sinal após a filtragem através do método BSS. Conclui-se que o ruído do sinal original foi extraído, notando que os picos estão mais salientes no sinal filtrado.

#### 4.1.2 Curtograma rápido e análise envelope

O primeiro método utilizado de pré-processamento foi uma função designada por curtograma rápido (fast kurtogram), associado a uma análise envelope, desenvolvido por Jerome Antoni (Jerome Antoni, 2015a)(Jerome Antoni, 2015b).

O curtograma rápido corresponde a um mapa formado pelo *Spectral Kurtosis* (SK) dividido em W janelas, em função da frequência f. É utilizado muitas vezes em conjunto com a análise envelope porque consegue determinar um filtro passa-banda ótimo para efetuar a análise envelope através do nível máximo de SK e de uma janela Wn, o qual fornece a frequência central e a largura de banda do passa-banda(Wang & Markert, 2015).

A função Matlab utilizada encontra-se na fig. 30. Através desta função, o sinal filtrado anteriormente é sujeito ao curtograma rápido e a análise envelope. Os resultados encontram-se nas fig. 31 e 33.

```
[cr,v,r,sk,sd,vpp, med, curtose, c] = Fast_Kurtogram(Sig_BSS_X,nlevel,Fs);
```

Figura 30-Função Matlab aplicada para realizar o curtograma rápido e a análise envelope ao sinal filtrado anteriormente

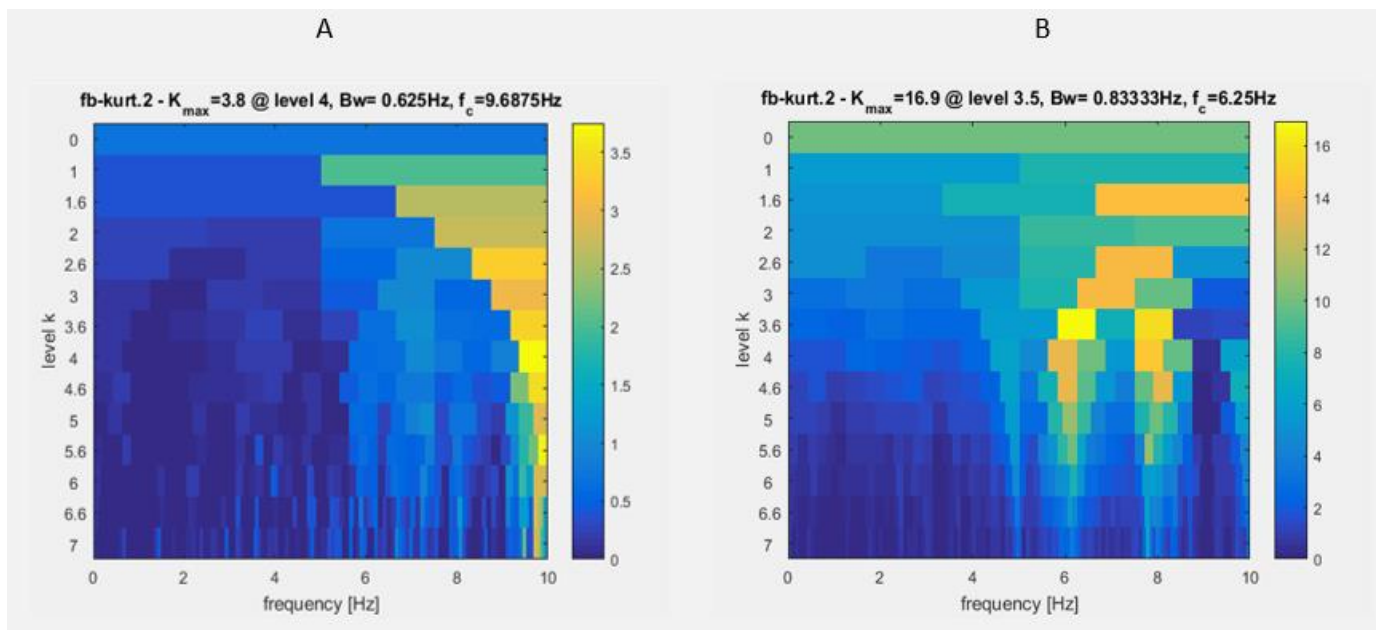


Figura 31-Rápido curtograma do sinal de vibração sem falha (A) e com falha (B).

Na fig. 31 está representado o curtograma rápido de dois sinais diferentes. Para um certo nível de decomposição, fornecido pelo utilizador (neste caso o nível foi 7), o curtograma vai decompor o sinal em frequência (fig. 32), analisando a curtose de cada janela. Valores altos significam grandes quantidades de picos e, conseqüentemente, mais probabilidade de se encontrarem frequências características das falhas do equipamento monitorizado.

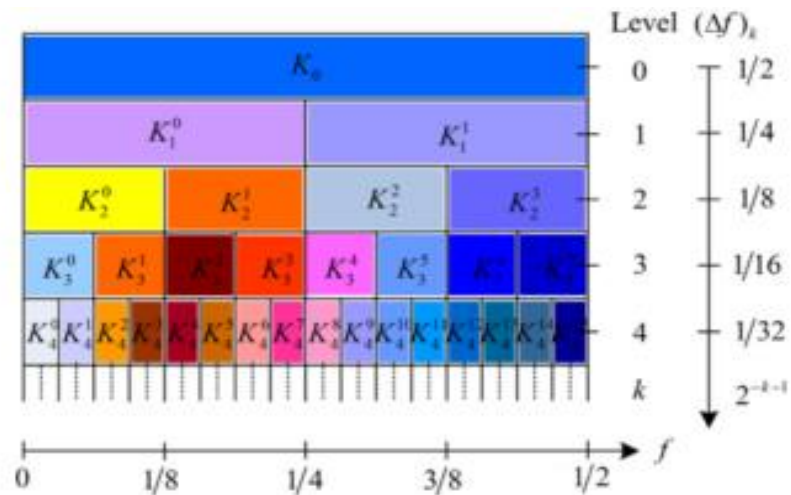


Figura 30-Árvore curtograma. Adaptado de [3]

A partir deste curtograma retiramos a zona ótima para analisar o sinal. Esta zona caracteriza-se então pelo maior nível de curtose, daí retiramos a frequência central e o nível de decomposição onde essa banda se encontra para realizar a análise envelope. Por exemplo, na fig. 31B (sinal com falha), a zona ideal é a zona a amarelo claro. Retiramos a frequência central de 6.25 Hz e o nível de decomposição onde se encontra, o nível 3.6, depois realiza-se a análise envelope.

Na fig. 33 está representada a análise envelope dos sinais da fig. 31. Analisando o sinal no domínio da frequência, e sabendo as frequências características das falhas descobriríamos as falhas ocorridas nos rolamentos.

Os artigos que utilizaram a análise envelope como método de diagnóstico de falha (Kamaras et al., 2016; Tian et al., 2016), calculam a frequência de falha característica dos rolamentos através das fórmulas da tabela 4.

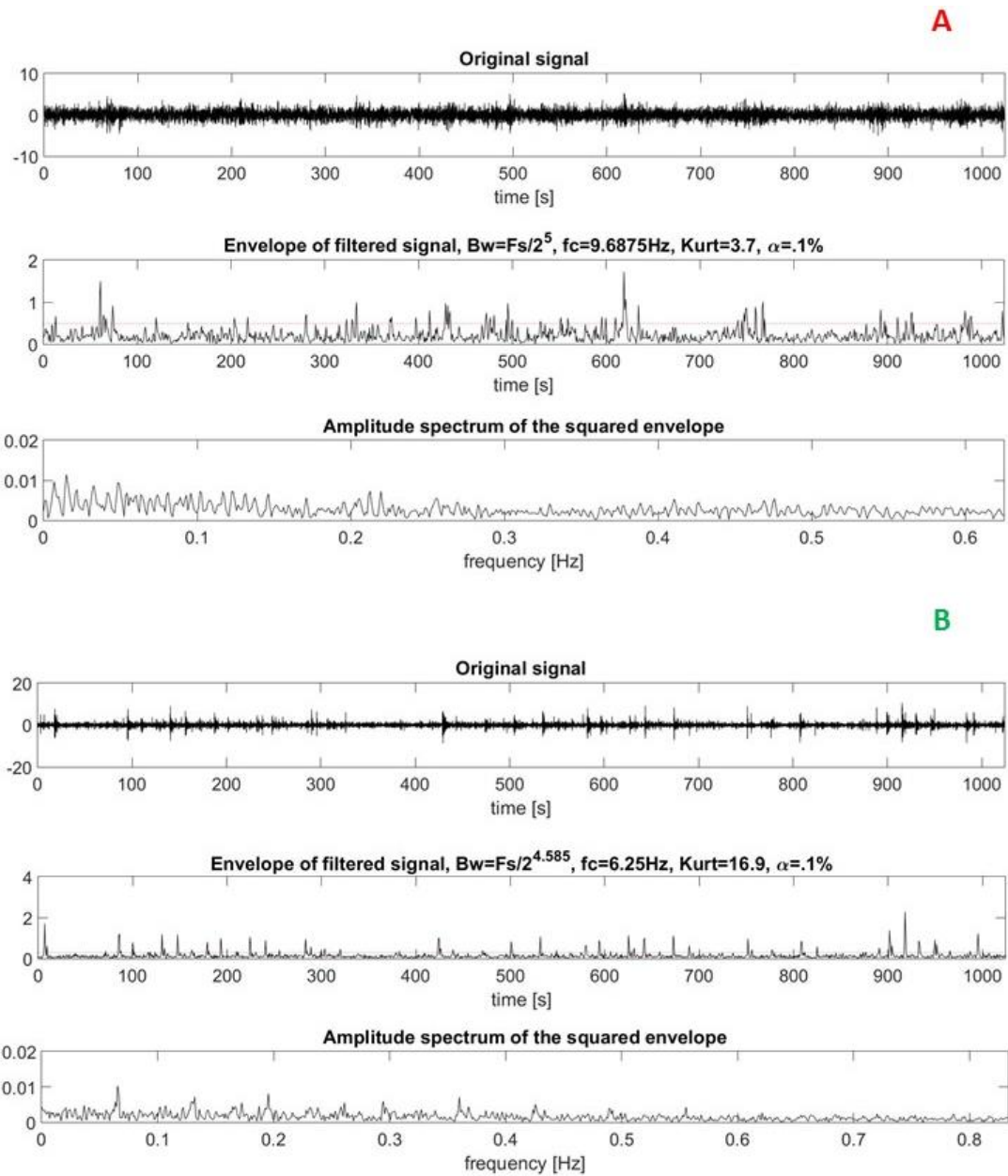


Figura 31-Análise envelope dos sinais sem falha (A) e com falha (B).

Tabela 4-Frequência de falha dos rolamentos.

|                                                                                                |                                                                     |
|------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|
| <p>Frequência da esfera passante na calha exterior (<i>Ball Pass Frequency Outer Race</i>)</p> | $f_s * \frac{Nb}{2} * \left(1 - \frac{Bd}{Pd} * \cos\varphi\right)$ |
|------------------------------------------------------------------------------------------------|---------------------------------------------------------------------|

|                                                                                                                                                                                |                                                                                                     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|
| Frequência da esfera passante na calha interior (Ball Pass Frequency Inner Race)                                                                                               | $f_s * \frac{Nb}{2} * \left(1 + \frac{Bd}{Pd} * \cos\varphi\right)$                                 |
| Frequência de rotação da esfera ( <i>Ball Spin Frequency</i> )                                                                                                                 | $f_s * \left(\frac{Pd}{2Bd}\right) * \left(1 - \left(\frac{Bd}{Pd}\right)^2 * \cos^2\varphi\right)$ |
| Frequência fundamental de rotação da gaiola ( <i>Fundamental Train Frequency</i> )                                                                                             | $f_s * \frac{1}{2} * \left(1 - \frac{Bd}{Pd} * \cos\varphi\right),$<br>para a calha exterior        |
|                                                                                                                                                                                | $f_s * \frac{1}{2} * \left(1 + \frac{Bd}{Pd} * \cos\varphi\right),$<br>para a calha interior        |
| <p>Fs – frequência de rotação</p> <p>Nb – número de esferas</p> <p>Bd – diâmetro da esfera</p> <p>Pd – diâmetro da gaiola</p> <p><math>\varphi</math> – ângulo de contacto</p> |                                                                                                     |

Devido ao facto de não se dispor de qualquer informação sobre os rolamentos em teste, não é possível calcular as suas frequências de falha. Analisando alguns sinais com e sem falha não se conseguiu estabelecer nenhuma semelhança ou encontrar algum padrão.

Considerando este método para realizar a monitorização num meio industrial real chega-se à conclusão que não é o mais indicado visto que, de máquina para máquina os parâmetros característicos dos elementos rotativos variam, por isso para utilizar este método seria necessário que a empresa tivesse todas as informações dos rolamentos constituintes da máquina, o que normalmente não acontece.

## 4.2 Wavelet Denoise + Wavelet Packet Transform

Com o método anterior não foi possível chegar a qualquer conclusão, o próximo método é completamente diferente, de forma a superar as limitações do procedimento anterior.

Para filtrar o sinal utilizou-se o método *Wavelet Denoise* (WD) e como pré-processamento foi utilizado a *Wavelet Packet Transform* (WPT). O diagrama do procedimento encontra-se na fig. 34.

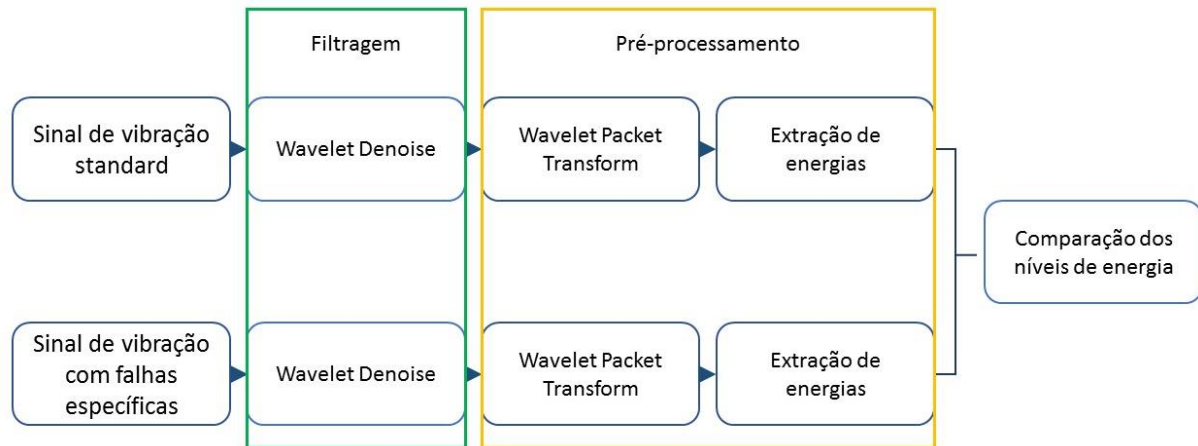


Figura 32-Diagrama da implementação WD + WPT

Neste procedimento utilizam-se dois conjuntos de dados diferentes, o conjunto de dados de referência e o conjunto de dados com falhas específicas discutido no **subcapítulo 3.1.3**.

O objetivo desta conduta é comparar as energias extraídas de cada conjunto de dados como acontece em (Zhigang, Hongmao, & Ning, 2016).

Como foi mencionado anteriormente, o método BSS não é adequado para filtrar sinais dos quais se pretende extrair características clássicas dependentes de amplitudes ou fase. Assim, implementou-se outro método de filtragem para extrair o ruído, o método *wavelet denoise* (WD) ou *wavelet thresholding* (MathWorks, 2017).

#### 4.2.1 *Wavelet Denoise*

Este método decompõe o sinal original até ao nível desejado, aplica um limite (*threshold*), o qual vai diminuir os coeficientes do sinal, removendo o ruído de baixa amplitude ou indesejado, depois reconstrói o sinal aplicando a transformada de wavelet inversa (Moshe, 2004).

Para aplicar esta técnica ao sinal de vibração, utilizou-se a função que se encontra na fig. 35. Introduziu-se o sinal a ser filtrado e foi escolhido um *soft threshold*, pois fornece resultados com curvas mais suaves, em comparação com o *hard threshold* (Moshe, 2004). O

nível de decomposição escolhido foi o nível 1 para que o sinal filtrado não perdesse muita informação do sinal original. A wavelet mãe utilizada foi a Daubechies 5.

```
New_X = wden(X, 'sqtwolog', 's', 'mln', 1, wavelet);
```

Figura 33- Função Wavelet Denoise

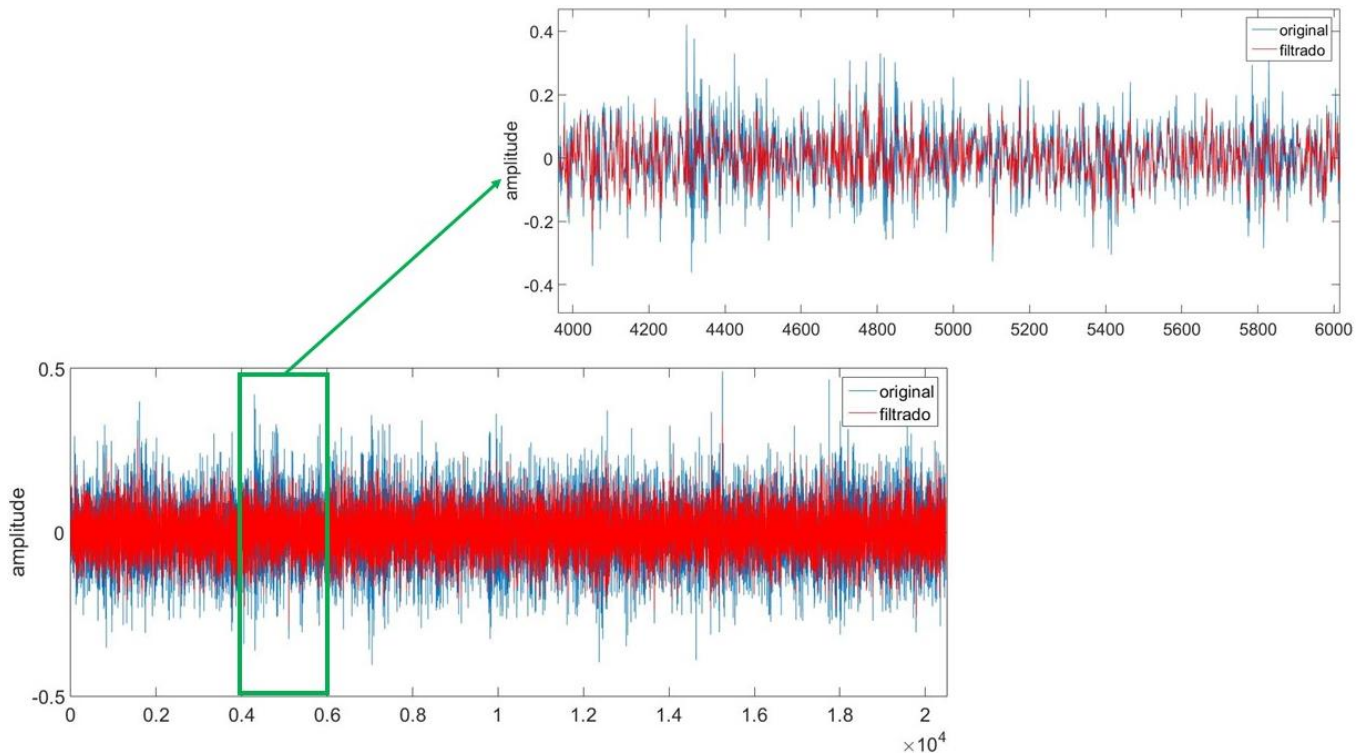


Figura 34-Denoise do sinal de vibração

Na fig. 36 está representado o sinal de vibração original a azul, e o sinal filtrado através do método WD, a vermelho. Observando a figura conclui-se que o ruído existente no sinal original diminuiu.

#### 4.2.2 Wavelet Packet Transform

Após a filtragem utilizou-se a WPT no pré-processamento dos dados.

A energia dos nós foi obtida com base em (Zhigang et al., 2016). Neste artigo é extraída a energia dos coeficientes (nós) da árvore *wavelet* de nível 3, que são utilizadas para classificar o sinal em diferentes tipos de falhas. Posto isto, retirou-se a energia do conjunto de sinais de referência e dos sinais com falhas específicas, usando a *wavelet* mãe Daubechies 8, e decompôs-se as amostras de vibração em 3 níveis.

Na fig. 35 está representada a função WPT, onde era decomposto o sinal (wpdec) e retiradas as energias dos coeficientes (wenergy).

A função de decomposição recebe o sinal filtrado, o nível de decomposição desejado e a wavelet mãe escolhida. Esta decomposição resultava na a árvore wavelet (T\_X) de onde, posteriormente são extraídas as energias.

```

1  function [T_X, Energy_X] = decomp_WPT(X,level_dec,wavelet)
2
3  T_X = wpdec(X,level_dec,wavelet);
4  % T_Y = wpdec(Y,level_dec,wavelet);
5
6  Energy_X = wenergy(T_X);
7  % Energy_Y = wenergy(T_Y);
8

```

Figura 35-Função de decomposição e extração da energia dos coeficientes do sinal, utilizando a WPT

Os resultados da comparação do nível de energia de cada nó extraído de ambos os sinais de vibração estão expostos nas figuras seguintes.

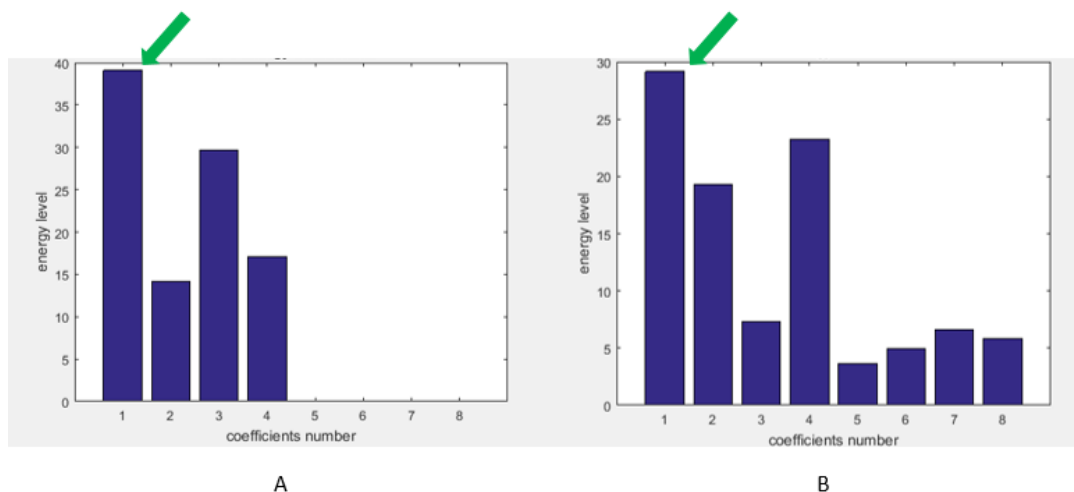


Figura 36-Energia dos coeficientes WPT de amostras sem falha (A-dados de referência, B-dados com falhas específicas)

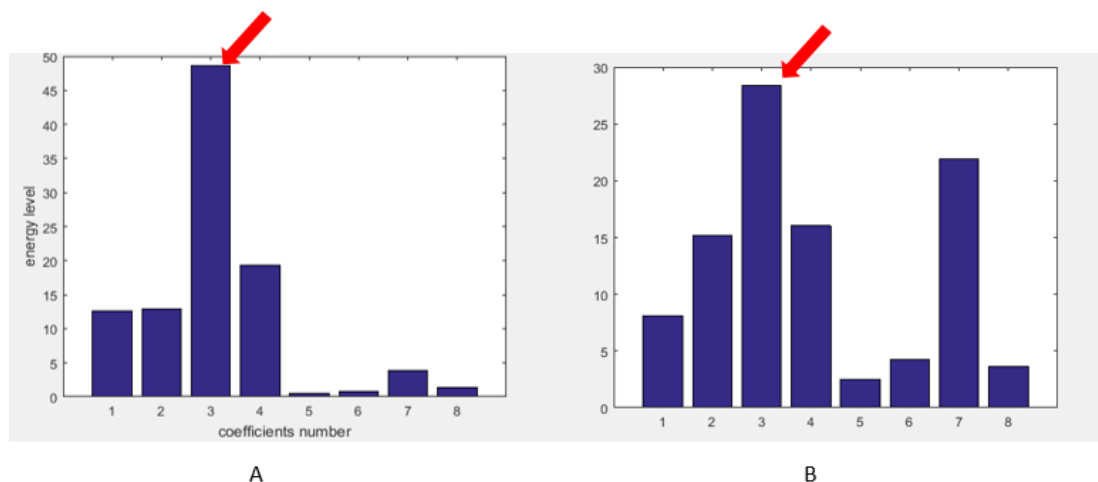


*Figura 37-Energia dos coeficientes WPT de amostras com falha (A-dados de referência, B-dados com falhas específicas)*

Na fig. 36 foram introduzidos os dados sem falha, e na fig. 37 foram testados os dados com falha na calha interior do rolamento. Para cada figura, o gráfico A corresponde ao conjunto de dados de referência e o gráfico B corresponde ao conjunto de dados com falhas específicas.

Nota-se que não existem muitas semelhanças, o que se deve ao facto de os sinais terem origem em máquinas diferentes. No entanto existem duas energias em cada um dos casos, falha e sem falha, que se destacam. No caso representado na fig. 36, o coeficiente com valor de energia mais elevado é o coeficiente número um. Em contraste, nas amostras com falha na calha interna do rolamento, o coeficiente número 3 é o que contém um valor de energia maior.

Devido a estes resultados, optou-se por calcular a energia de amostras do grupo de dados de referência ao longo do tempo. Os resultados estão na fig. 38.

Analisando os vários gráficos, chega-se à conclusão que a energia do coeficiente número 1 vai diminuindo à medida que a energia do coeficiente número 3 vai aumentando consideravelmente ao longo do tempo.

Posto isto, conclui-se que a energia é uma boa técnica para identificar vários tipos de falha. No entanto, é preciso que existam dados de vibração de cada tipo de falha provenientes da mesma máquina a monitorizar. Diferentes tipos de falha manifestam-se em diferentes coeficientes, permitindo assim detetar e prever falhas.

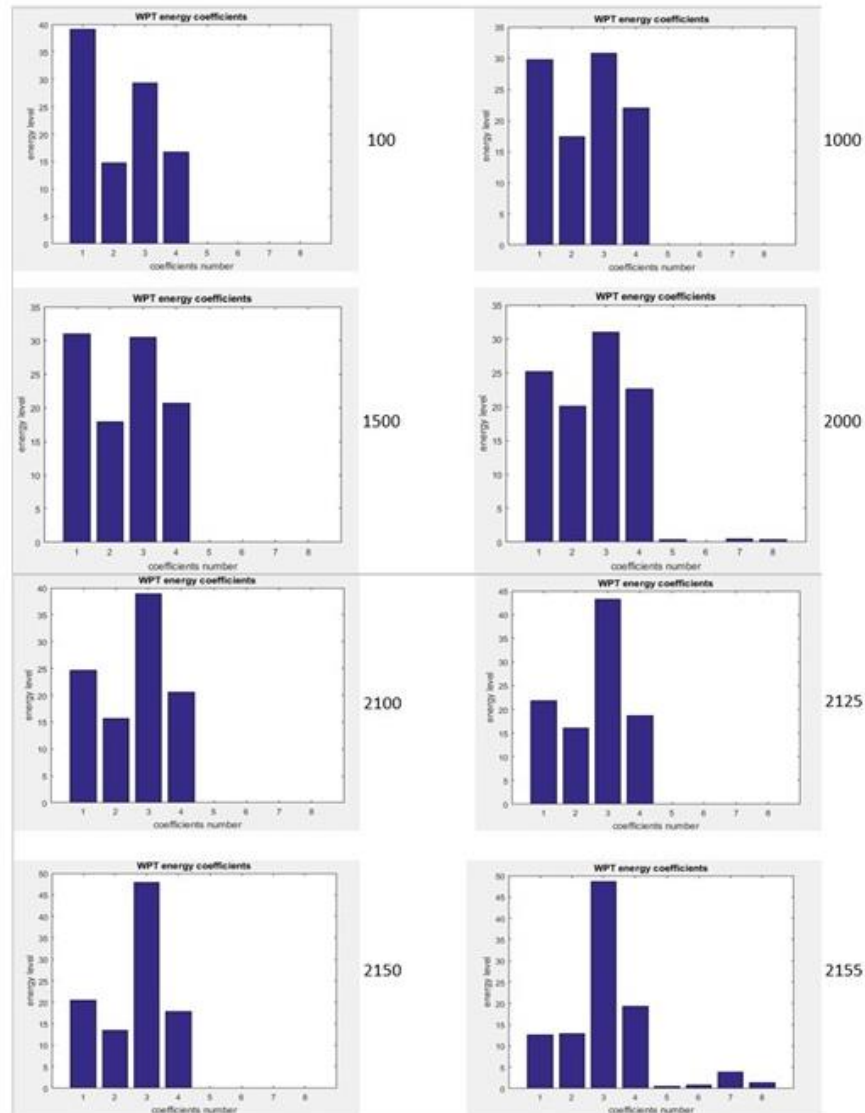


Figura 38-Histórico dos coeficientes de energia do conjunto de amostras de referência

### 4.3 Conclusão

Neste capítulo foram apresentadas duas implementações distintas inseridas numa abordagem estatística.

Na primeira implementação recorreu-se ao método BSS para fazer a filtragem dos sinais e à união do curtograma rápido com a análise envelope para fazer o seu pré-processamento e deteção de falhas.

Esta implementação era promissora, no entanto, sem as frequências características dos rolamentos não é possível detetar as falhas através da análise envelope. Chegou-se ainda à conclusão que não era um método adequado para realizar a monitorização no meio industrial, pois a informação sobre os rolamentos das máquinas não se encontra sempre disponível, e assim, não se consegue calcular as frequências características das diferentes falhas.

Na segunda implementação para filtragem dos dois conjuntos de dados utilizou-se o método WD e a WPT para pré-processamento.

Esta técnica mostrou bons resultados, pois através dos coeficientes dos sinais de vibração conseguiu-se detetar a evolução da falha. No entanto, necessita-se de dados de vibração característicos de cada tipo de falha para cada máquina a monitorizar, o que nem sempre está disponível.

Como nenhuma das implementações era adequada para resolver o problema em questão, decidiu-se implementar novas técnicas de processamento de sinal e de extração e aprendizagem do sistema.

No próximo capítulo vão ser introduzidos métodos adequados para transformar as características extraídas em informação útil através da aprendizagem supervisionada.

## 5. ABORDAGEM DE APRENDIZAGEM SUPERVISIONADA DO SISTEMA

Devido aos resultados insatisfatórios da abordagem estatística, decidiu-se implementar algoritmos baseados em aprendizagem supervisionada para categorizar a informação extraída e realizar a aprendizagem do sistema.

Para realizar o pré-processamento utilizou-se a DWT. Os algoritmos de aprendizagem supervisionada utilizados foram a SVM e o FCM. Na filtragem foi utilizado o método WD discutido no capítulo anterior e a normalização utilizada é a normalização standard.

No desenvolvimento, o grupo de dados utilizado inicialmente foi o grupo de dados de referência com o objetivo de testar a performance das combinações de algoritmos. Nas implementações que produziram bons resultados, foram posteriormente introduzidos os dados cedidos pela empresa.

### 5.1 Discrete Wavelet Transform + Support Vector Machine

A técnica de pré-processamento DWT é uma transformada *wavelet* muito utilizada na extração de características de rolamentos.

A SVM trata-se de um classificador binário supervisionado muito utilizado na análise de dados de vibração.

Devido às suas características estes dois métodos foram implementados em conjunto. O diagrama do procedimento encontra-se na fig. 39.

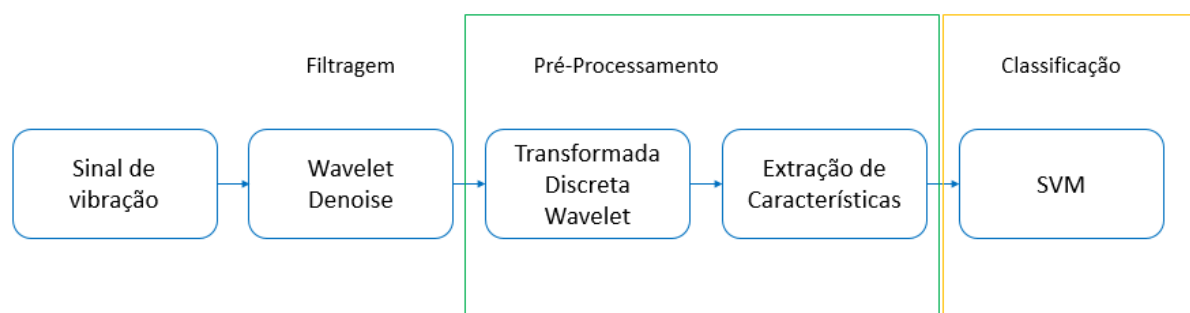


Figura 39-Diagrama de implementação DWT + SVM

### 5.1.1 Discrete Wavelet Transform

No processamento das amostras de vibração utilizando a DWT foi sempre utilizado um nível de decomposição 5 e com uma *wavelet* mãe Daubechies 8.

Os parâmetros das *wavelets* foram escolhidos com base em (Vujovi, 2015), que indica que as melhores *wavelets* mãe são Daubechies 7 e 8 e o melhor nível de decomposição é o nível 5.

Utilizando a DWT foram extraídas características estatísticas (secção 3.3.2) de cada amostra de vibração, e a seguir introduzidas na SVM. A função DWT utilizada e a extração das características estatísticas está representada na fig. 40.

```
function [RMS_X, STD_X, Energy_X, MED_X, VPP_X, K_X, SK_X, VAR_X, CRESTF_X] = decomp_DWT(X,level_dec,wavelet)
    %%%%%%%Características do sinal no eixo dos XXs%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

    [coefX,~] = wavedec(X,level_dec,wavelet);

    RMS_X = rms(coefX);
    STD_X = std(coefX);
    Energy_X = sum(coefX.^2); %energia dos coeficientes %C_AP.^2 faz a potencia de 2 a cada elemento de matri:
    MED_X = mean(coefX);
    VPP_X = peak2peak(coefX);
    K_X = kurtosis(coefX);
    SK_X = skewness(coefX);
    VAR_X = var(coefX);
    maxvalue = max(coefX);
    CRESTF_X = abs(maxvalue)/RMS_X;
```

Figura 40-Função de pré-processamento DWT e extração de características estatísticas

### 5.1.2 Support Vector Machine

Quando se utilizam classificadores como a SVM, o conjunto de dados que vai ser analisado tem que ser dividido em dois grupos de dados, o conjunto de treino, que permite que o sistema aprenda, e o conjunto de teste, que testa a eficácia do classificador na identificação de padrões e falhas. Um exemplo de funções de treino e teste da SVM estão representadas na fig. 41.

Neste tipo de classificação é importante que o conjunto de treino seja representativo do sinal, ou seja, tem que conter dados com e sem falha, para que o classificador consiga realizar uma boa classificação.

```
SVMMModel = fitcsvm(m,ind,'KernelFunction','rbf','Standardize',true);
[L,s] = predict(SVMMModel,M2);
```

Figura 41-Funções de treino e teste da SVM

O conjunto de dados de referência foi dividido em 25% do total de amostras para treino e 75% para teste. O conjunto de treino era constituído por 20% das amostras que caracterizam o bom funcionamento da máquina (amostras iniciais), e 5% de amostras com falha (amostras finais).

Como não se sabe onde ocorre a falha ao longo das amostras, esta separação é realizada de uma forma hipotética.

Os testes foram realizados de dois modos diferentes: um em que é introduzido na SVM o conjunto de treino (25%) seguido do conjunto de teste (75%); outro teste apenas é introduzido o conjunto de teste (75%).

Estes dois testes foram realizados para verificar se a classificação em que o conjunto de teste introduzido era composto pelo conjunto total de amostras era melhor que a classificação só com o conjunto de teste.

Para facilitar a interpretação das figuras correspondentes à classificação da SVM foi obtida a porção de amostras classificadas como falha ao longo do conjunto de dados (fig. 44).

Estes gráficos foram construídos separando o número de amostras em vetores de 300 amostras (quando no teste é utilizado o conjunto de treino) ou em 200 (quando no teste é só utilizado o conjunto de teste) e depois, em cada vetor, são contabilizadas as amostras classificadas como falha, como se pode ver na fig. 42.

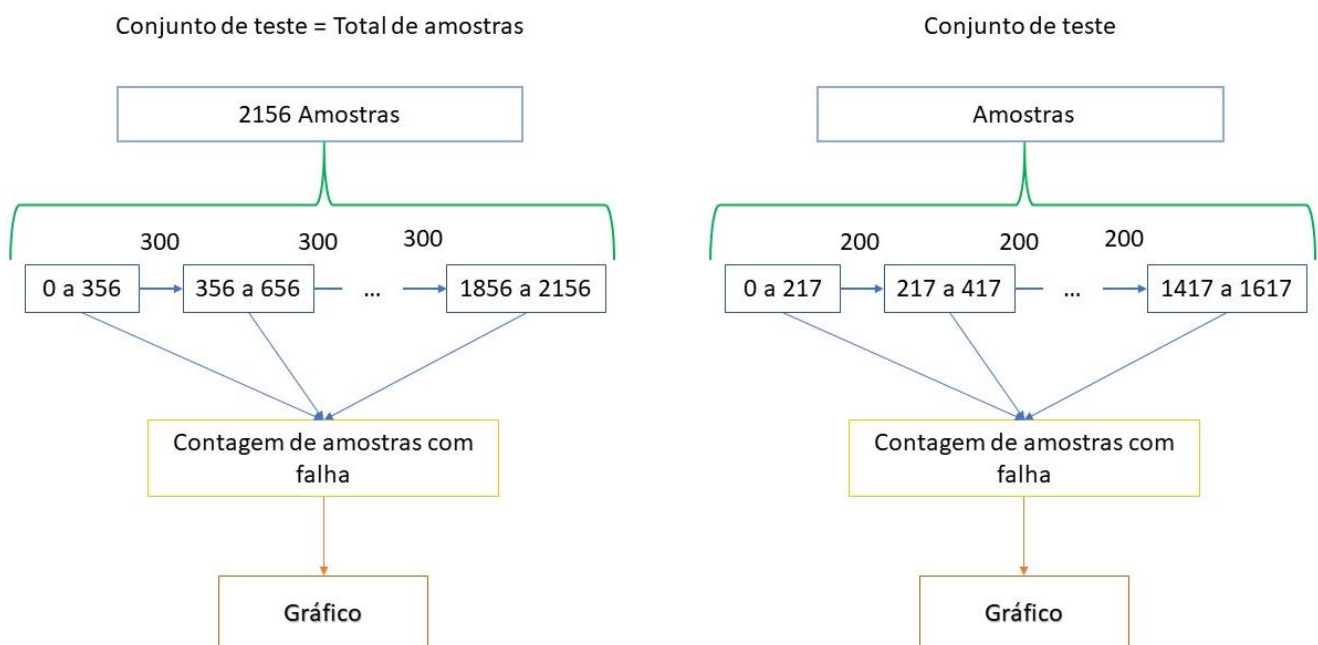


Figura 42-Procedimento de contagem de amostras com falha para os dois testes realizados

A classificação realizada pela SVM pode ser vista nas figuras 43 e 44. Às amostras sem falha corresponde a classe 0 e às amostras com falha corresponde a classe 1.

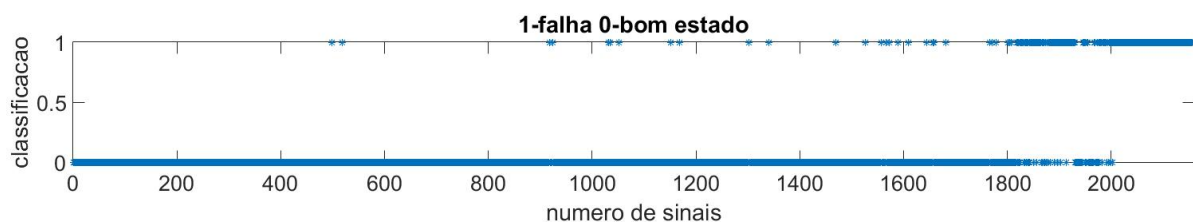


Figura 43-Classificação da SVM com conjunto de total de amostras (conjunto de treino + conjunto de teste)

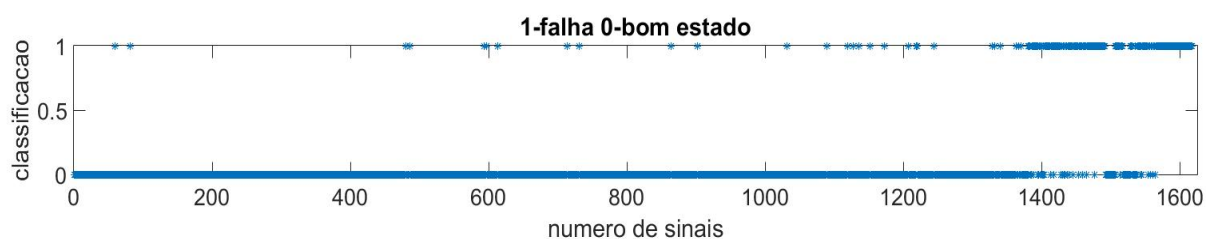


Figura 44-Classificação da SVM com conjunto de teste

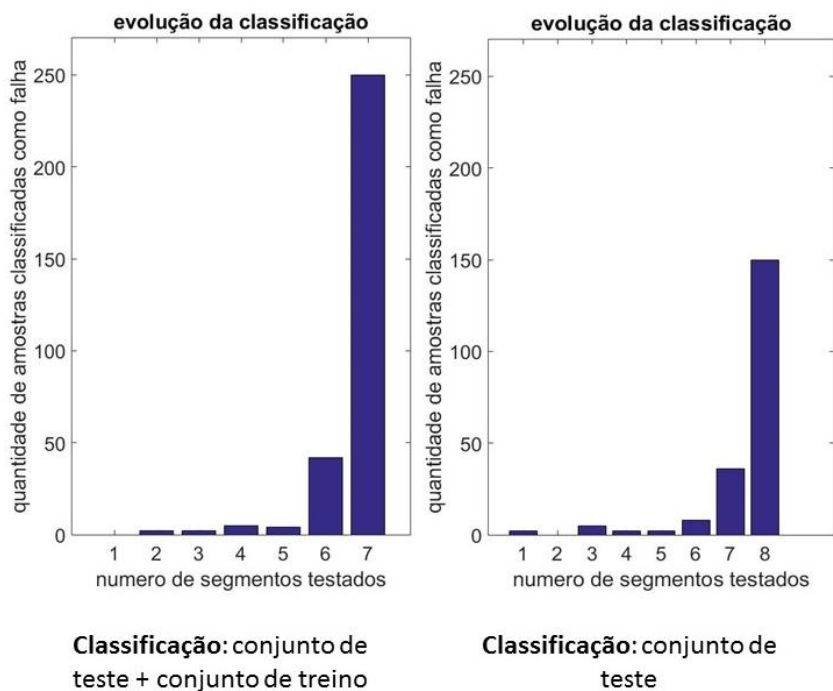


Figura 45- Quantidade de amostras classificadas como falha nos gráficos da fig. 42 e 43 respetivamente

| Sem dados de treino |        |    |         |      |         | Com dados de treino |        |    |        |      |         |
|---------------------|--------|----|---------|------|---------|---------------------|--------|----|--------|------|---------|
| 1                   | 1.1872 | 43 | 1.3966  | 2093 | -0.2034 | 1                   | 1.1221 | 43 | 1.1961 | 2093 | -0.0692 |
| 2                   | 0.8741 | 44 | 0.9559  | 2094 | 0.0091  | 2                   | 1.0003 | 44 | 1.0560 | 2094 | 0.8889  |
| 3                   | 1.2150 | 45 | 0.7591  | 2095 | -0.5623 | 3                   | 1.4864 | 45 | 1.2698 | 2095 | -0.7160 |
| 4                   | 0.7635 | 46 | 1.5555  | 2096 | -0.8355 | 4                   | 1.4782 | 46 | 1.4911 | 2096 | 0.6922  |
| 5                   | 1.2975 | 47 | 1.5770  | 2097 | -0.7698 | 5                   | 1.2772 | 47 | 1.2689 | 2097 | -0.6527 |
| 6                   | 1.0560 | 48 | -0.0764 | 2098 | -0.6157 | 6                   | 1.0003 | 48 | 1.1583 | 2098 | -0.7313 |
| 7                   | 0.8053 | 49 | 0.9219  | 2099 | -0.7052 | 7                   | 1.0001 | 49 | 1.2959 | 2099 | 0.1577  |
| 8                   | 1.0559 | 50 | 1.2947  | 2100 | -0.7671 | 8                   | 1.5750 | 50 | 1.3165 | 2100 | -0.1541 |
| 9                   | 1.3031 | 51 | 1.1128  | 2101 | -0.9942 | 9                   | 1.3429 | 51 | 0.1404 | 2101 | 0.5332  |
| 10                  | 1.2860 | 52 | 1.2803  | 2102 | -0.7748 | 10                  | 1.2954 | 52 | 1.2405 | 2102 | 0.1473  |
| 11                  | 1.3244 | 53 | 0.9065  | 2103 | -0.9261 | 11                  | 0.9372 | 53 | 1.3989 | 2103 | -0.3205 |
| 12                  | 1.2922 | 54 | 1.1951  | 2104 | -0.9254 | 12                  | 1.0004 | 54 | 1.4488 | 2104 | -0.3826 |

Figura 46-Alguns resultados do vetor de probabilidades da classificação da SVM sem e com conjunto de treino

As figuras 43 e 44 apresentam a classificação realizada pela SVM. Embora a tendência da classificação seja clara, para facilitar a interpretação das figuras, nos gráficos da fig. 45 é exposta a quantidade de amostras classificadas como falha ao longo do conjunto de dados.

Como foi referido no início do capítulo, os dados utilizados nesta implementação correspondem aos dados de referência. É sabido que este conjunto de sinais possui uma falha no final do seu conjunto, não se sabendo exatamente onde se encontra. Assim, observando o gráfico obtido através da SVM (fig. 43 e 44), conclui-se que o classificador conseguiu identificar a falha. Inicialmente há poucas evidências de falha, mas esta torna-se clara nas amostras finais do conjunto de dados, onde a classe 1 aumenta. Este resultado está em concordância com a informação do conjunto de dados utilizado.

Na fig. 46, o vetor referente às probabilidades de classificação indica-nos o grau de confiança com que foi realizada a classificação de cada amostra. A SVM classifica as amostras entre 1 e -1. Quando maior for a aproximação da probabilidade a estes valores maior certeza há na classificação da respetiva amostra. Pode então concluir-se que o algoritmo conseguiu identificar as amostras com uma probabilidade razoável.

Embora se tenha conseguido resultados satisfatórios com este método e utilizando este conjunto de amostras, não se consegue classificar outros conjuntos de teste corretamente utilizando o mesmo conjunto de treino, como por exemplo, utilizar o conjunto de treino de referência e como conjunto de teste os dados fornecidos pela empresa.

Para além do classificador não ser adaptativo, a principal razão para não se utilizar conjuntos de dados de treino e teste provenientes de máquinas diferentes, é o facto de as

máquinas de onde foram extraídos os dois grupos de dados utilizados serem diferentes, logo as máquinas não têm a mesma “assinatura” de falha. Assim, tem que ser dado ao algoritmo um novo conjunto de treino característico do novo caso, para que ele o possa aprender e identificar no conjunto de teste.

Como se obtiveram classificações razoáveis, testou-se este procedimento com os dados cedidos pela empresa Controlar. O método de separação deste conjunto de amostras em conjunto de treino e teste foi o mesmo utilizado para o conjunto amostras de referência. Os resultados encontram-se nas figuras seguintes.

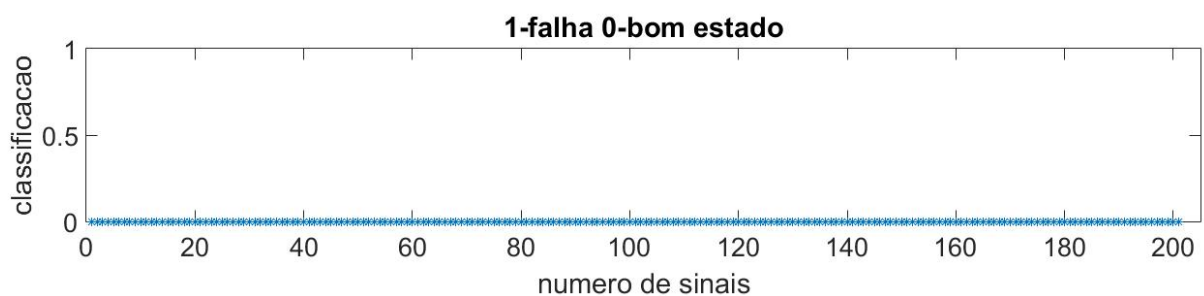


Figura 47- Classificação da SVM com conjunto de total de amostras (conjunto de treino + conjunto de teste)

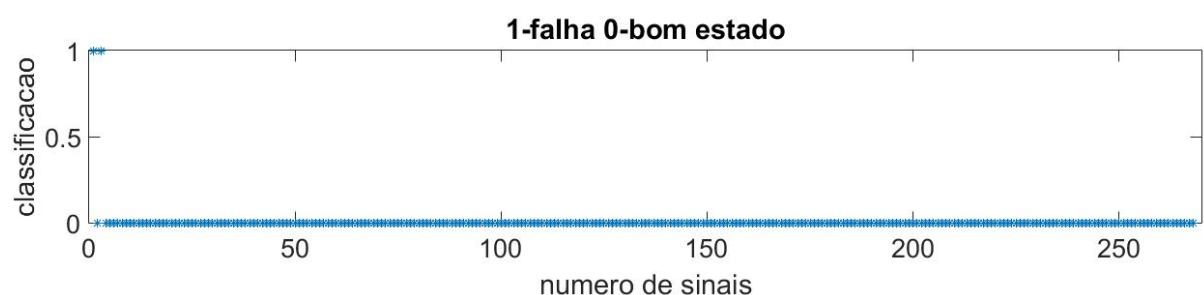


Figura 48-Classificação da SVM com conjunto de teste

Como se pode observar nas fig. 47 e 48, os resultados não são os melhores. O conjunto de treino não é representativo do sinal, no procedimento de aprendizagem da SVM, esta foi treinada com mais dados sem falha do que com falha e, como não é um algoritmo adaptativo não conseguiu classificar as amostras adequadamente.

## 5.2 Discrete Wavelet Transform + Fuzzy C-Means Clustering

Na tentativa de melhorar a classificação anterior, decidiu-se utilizar na classificação um algoritmo supervisionado e adaptativo, o FCM.

Para testar esta implementação foram apenas utilizados os dados de referência. Os dados foram normalizados com a normalização standard.

### 5.2.1 Discrete Wavelet Transform

Utilizando a DWT para pré-processar os dados, como foi realizado em 5.1.1, são extraídas 8 características estatísticas de todas as amostras de vibração. Estas características foram introduzidas diretamente no algoritmo FCM.

O diagrama de implementação está na fig. 49.

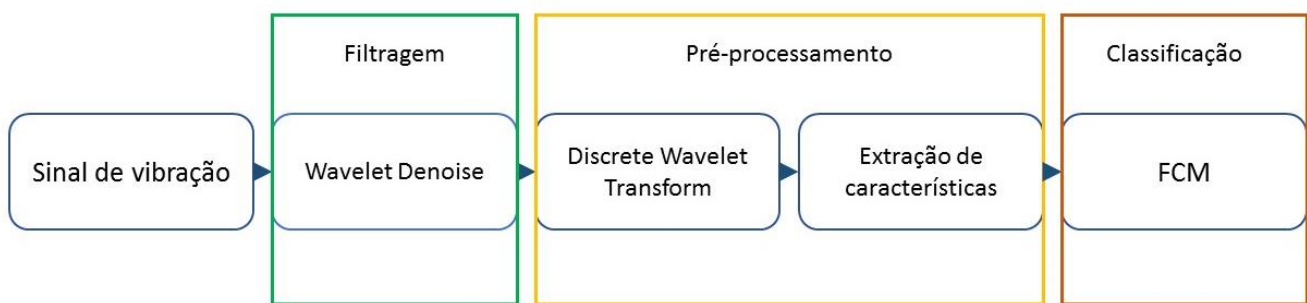


Figura 49-Diagrama de implementação DWT + FCM

### 5.2.2 Fuzzy C-Means Clustering

Após a introdução das características estatísticas no algoritmo, utilizando a função da fig. 50, os seus resultados podem ser vistos nas fig. 51 e 52.

```
[centers, U] = fcm(m, 2, options);
```

Figura 50-Função Matlab do algoritmo Fuzzy C-Means Clustering

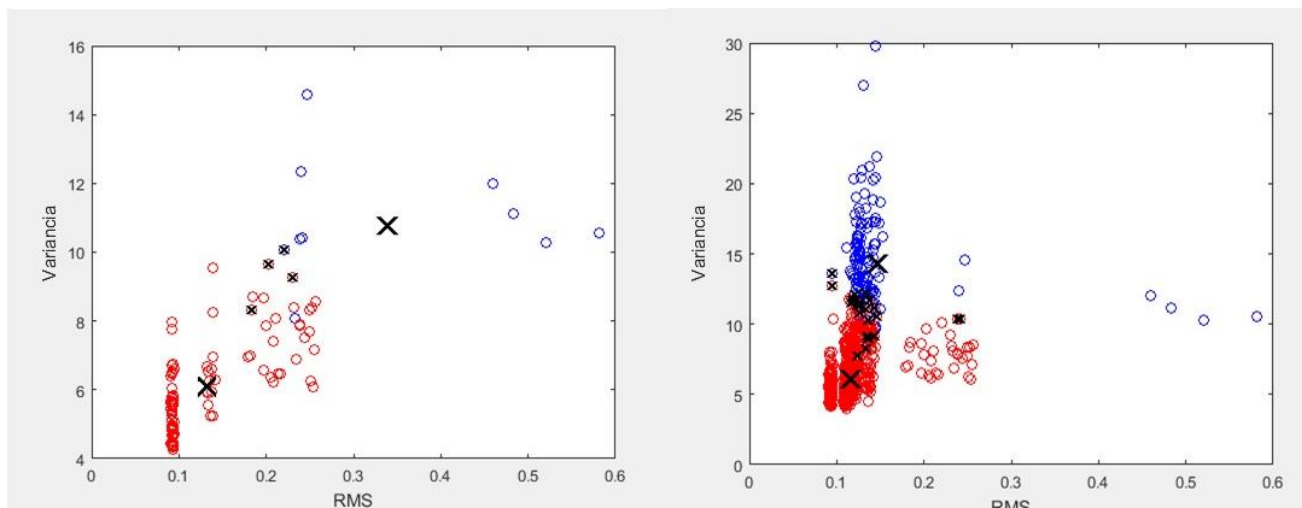


Figura 51-Classificação FCM para: A) 100 amostras e B) 1000 amostras

Analisando a fig. 51, observa-se a separação do conjunto de dados em dois clusters. Os X (x em maiúscula) representam os centros dos clusters e as amostras marcadas com x (x em minúscula) são as que têm menor grau de associação em relação aos dois clusters (amostras com maior incerteza de classificação). Aqui estão apenas retratadas duas características, o valor eficaz e a variância, pois eram as duas características onde a separação dos clusters era mais perceptível.

O algoritmo tem ainda como saída o grau de associação de cada amostra aos clusters existentes. Na fig. 52 está retratado isso mesmo relativamente ao cluster 1 (a vermelho) e ao cluster 2 (a azul).

Como se pode ver na fig. 52, o algoritmo com poucas amostras consegue diferenciar bem os dois clusters. À medida que são acrescentadas mais amostras e a falha no final do teste vai ficando cada vez mais evidente, o algoritmo piora a classificação, não conseguindo separar os dados com confiança.

Como não se sabe onde se encontra o início da falha é difícil dividir os dados para formar um bom conjunto de treino. Assim um classificador que, através da deteção de padrões e de relações dos valores de entrada, consiga detetar as falhas sem conhecer as respostas *à priori* é ideal para realizar uma boa classificação.

Além disso este algoritmo é muito sensível a parâmetros essenciais, puramente hipotéticos fornecidos pelo utilizador (como por exemplo, o número de clusters) influenciando a classificação.

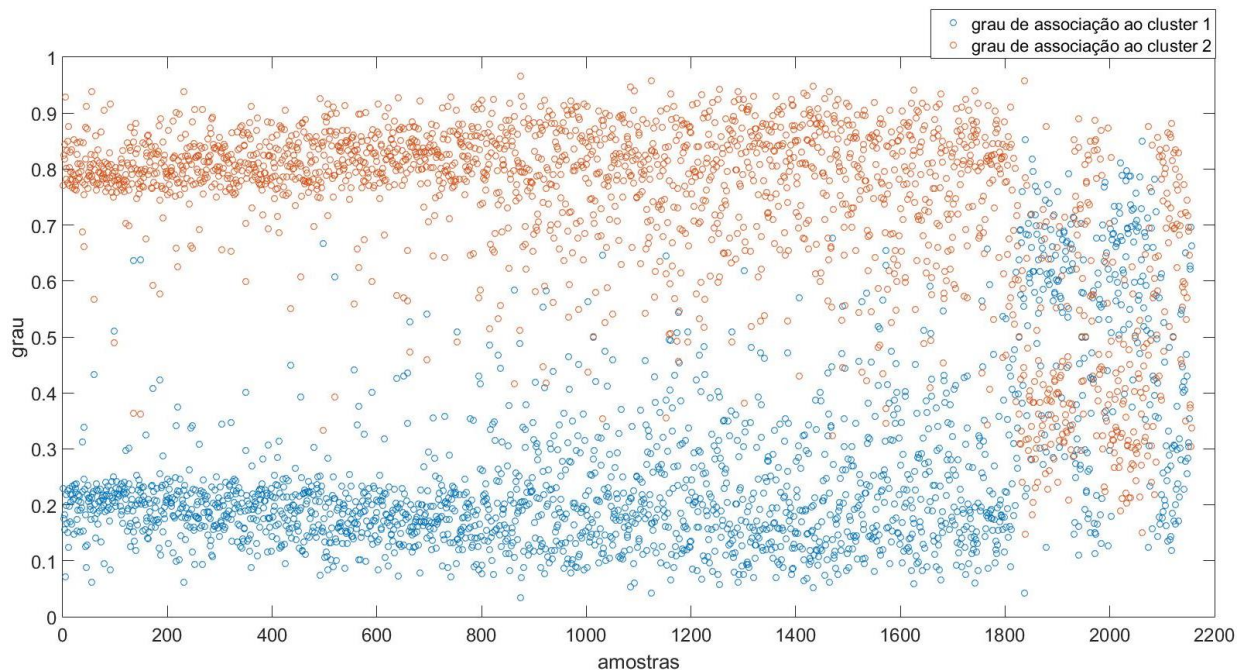


*Figura 52-Grau de associação de cada amostra aos dois clusters*

### 5.3 Conclusão

No presente capítulo foram implementados dois métodos diferentes recorrendo a uma abordagem de aprendizagem supervisionada.

Em ambos os métodos foi empregue a mesma técnica de pré-processamento de dados, a DWT, para extrair as características estatísticas das amostras de vibração.

Após o processamento, foi inicialmente desenvolvido um classificador binário supervisionado e não adaptativo muito utilizado na análise de dados de vibração, a SVM.

Embora, utilizando os dados de referência, a classificação tenha sido satisfatória, ao introduzir os dados de teste cedidos pela empresa com os dados de treino de referência não houve bons resultados, pois os dados provêm de máquinas diferentes e, assim, a “assinatura” da falha é diferente.

Quando foram introduzidos os dados cedidos pela empresa, como conjunto de treino e teste, a classificação foi igualmente má. O conjunto de treino possuía poucas amostras de falha comparativamente às amostras de bom funcionamento e, como o algoritmo não é

adaptativo, partindo das amostras de falha que tinha não foi capaz de identificar falhas iguais às do conjunto de treino.

Por essa razão, decidiu-se implementar um classificador adaptativo baseado em lógica *fuzzy*, o FCM.

As características estatísticas extraídas das amostras de referência foram introduzidas no algoritmo FCM, chegando-se à conclusão que, embora seja possível ver a progressão da falha, o classificador no suposto lugar da falha teve dificuldade em identificá-la com certeza, pois o seu grau de confiança na classificação das amostras respetivas é contestável.

Assim, decidiu-se implementar um algoritmo que recorresse à aprendizagem não supervisionada, para que através da relação de padrões existentes, consiga identificar, com mais confiança as falhas.

## 6. ABORDAGEM DE APRENDIZAGEM NÃO SUPERVISIONADA

Os algoritmos não supervisionados, não têm conhecimento *à priori* dos dados, tratando-os como objetos aleatórios, e com base nas relações e padrões descobertos constituem modelos.

Este tipo de aprendizagem é muito útil pois, a maior parte das vezes o utilizador não possui um conjunto de dados diversificado, constituído tanto por amostras que caracterizam o bom estado da máquina, como por amostras que assinalam as falhas do equipamento, essenciais ao treino de um classificador baseado em aprendizagem supervisionada, como a SVM.

No pré-processamento foi utilizado, para além da DWT, o método EMD. A normalização foi realizada utilizando o método descrito na secção 3.3.2. e foram utilizados dois classificadores não supervisionados, o OCSVM e a SVDD.

Mais uma vez, para testar a performance dos métodos implementados, inicialmente foram introduzidas as amostras de referência e posteriormente as amostras cedidas pela empresa Controlar.

### 6.1 Discrete Wavelet Transform/Empirical Mode Decomposition

A DWT e o EMD foram os dois métodos utilizados no pré-processamento dos conjuntos de dados utilizados.

Com a DWT, utilizando o nível de decomposição 7 e duas *wavelets* mãe distintas a Daubechies 8 e a Haar, extraíram-se as 8 energias (7 coeficientes de detalhe + 1 coeficiente de aproximação) de cada amostra.

A utilização da *wavelet* mãe Haar foi realizada para verificar se eram conseguidos melhores resultados, porque no estudo do artigo (M. et al., 2012), os autores conseguem melhores resultados ao utilizarem a referida *wavelet* mãe.

O método EMD, como já foi mencionado na secção 2.1.3, tem como objetivo a decomposição do sinal em monocomponentes (IMFs) o que lhe permite uma boa capacidade de adaptação e o torna altamente eficaz no tratamento de dados não estacionários (M. et al., 2012), como são os dados de vibração.

Este algoritmo foi implementado devido aos bons resultados alcançados em (M. et al., 2012). Com base nesse trabalho, o sinal de vibração foi decomposto em 2 e 8 IMFs e, as amostras de vibração foram também decompostas em 2 e 8 IMFs.

## **6.2 One-Class Support Vector Machine e Support Vector Data Description**

O OCSVM e SVDD são métodos adaptativos e não supervisionados, que precisam apenas de uma classe de dados para realizar a classificação. Isto torna-os adequados para a análise e detecção de falhas de operação de uma máquina industrial, pois no início desta análise, normalmente está apenas disponível um tipo (classe) de dados, geralmente dados que refletem o bom funcionamento da máquina (M. et al., 2012).

Estes métodos foram implementados pelo facto de serem extremamente adequados para monitorização de falhas nas máquinas industriais, e também, devido à falta de amostras que caracterizam falhas ocorridas nos rolamentos, necessários para técnicas de classificação de duas ou mais classes, como o caso da SVM.

A implementação destes algoritmos consiste em três fases: a primeira é o mapeamento dos dados num espaço de características de grande dimensão; a segunda é a escolha do kernel e cálculo da margem; e por fim, a terceira é a classificação para detecção de *outliers* (Joon et al., 2005).

Para implementar os dois métodos recorreu-se a uma biblioteca em desenvolvimento, a LIBSVM (C. Chang & Lin, 2011). Esta biblioteca permite realizar a classificação e regressão recorrendo a algoritmos baseados na SVM (C.-C. Chang & Lin, 2016), e está disponível para várias linguagens de programação, incluindo Matlab. É uma das ferramentas de software mais utilizadas para a implementação deste tipo de técnicas, pois é extremamente simples, eficiente (C. Chang & Lin, 2011) e versátil.

As características extraídas são representadas utilizando a representação *sparse*, fig. 53, pelo simples facto de que a biblioteca utilizada, a LIBSVM, apenas suporta este formato (C.-C. Chang & Lin, 2015). A matriz *sparse*, contém a matriz das características das amostras de vibração, e a matriz das etiquetas que representam a natureza dessas características.

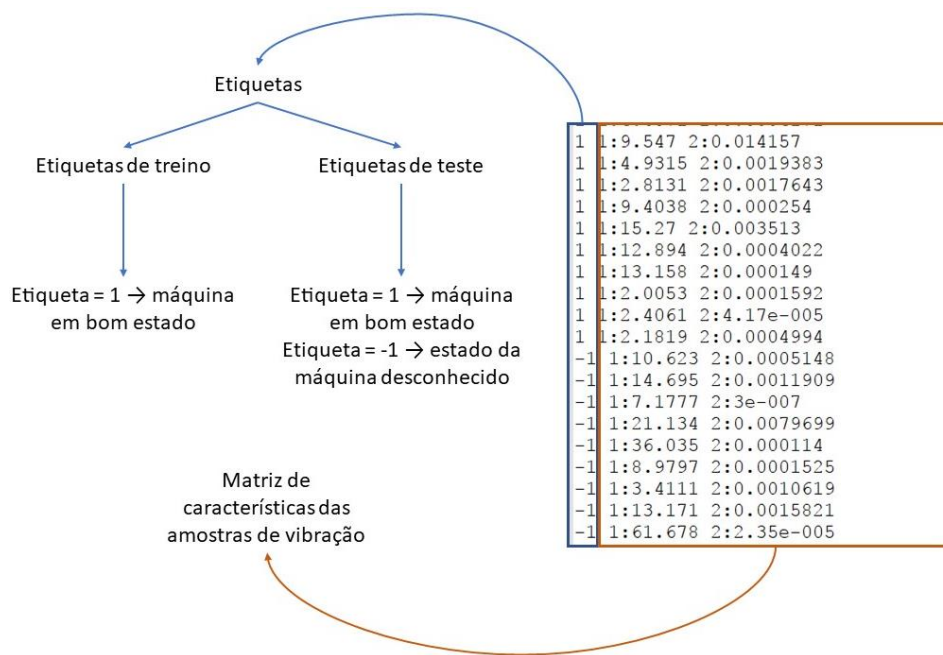


Figura 53-Representação das características extraídas numa matriz sparse

Para testar a performance do algoritmo, o conjunto de teste é constituído por características que representam o bom funcionamento (etiqueta 1) e o estado desconhecido da máquina (etiqueta -1).

Na implementação destes algoritmos é necessário determinar um kernel adequado, pois estas técnicas são muito sensíveis a eles (Joon et al., 2005). Em relação ao kernel, existem vários tipos e até existe a possibilidade de o utilizador desenvolver um a seu gosto. A escolha de um bom kernel é essencial para obter uma boa classificação, pois permite separar linearmente os dados que não são linearmente separáveis, mapeando-os num novo espaço de características, sem aumentar o custo computacional. Os melhores e mais utilizados são os kernels Gaussiano e o polinomial (Deák et al., 2014).

O kernel Gaussiano foi o escolhido tendo como base uma experiência realizada em (Joon et al., 2005), que indica que este kernel teve a melhor performance comparando com outros. Em (M. et al., 2012) também é feita essa comparação, chegando à mesma conclusão.

No kernel existem ainda parâmetros a determinar, que otimizam a classificação do algoritmo. Os melhores parâmetros são encontrados através do método de validação cruzada.

### 6.2.1 DWT/EMD + OCSVM

Para testar o método implementado, dividiu-se o conjunto de dados de referência da seguinte forma: 75% dos dados em boas condições como conjunto de treino e o resto dos dados como conjunto de teste.

A OCSVM apenas necessita de uma classe de dados para realizar o treino. Assim como as amostras que representam o bom funcionamento da máquina estão em maior quantidade, relativamente às amostras que representam falhas, decidiu-se utiliza-los no conjunto de treino. Além disso, num ambiente industrial, é mais comum haver dados de bom funcionamento da máquina do que dados de falha.

Como não se sabe onde começa exatamente a surgir a falha do rolamento, apenas se sabe que ela ocorre no final da experiência de teste-falha, dividiram-se os dados de forma hipotética, considerando as primeiras 1700 amostras, das 2156, como amostras que refletem a boa condição do rolamento.

No conjunto de dados cedidos pela empresa Controlar, utilizou-se como conjunto de treino todas as amostras de que se tinha conhecimento caracterizarem o bom funcionamento da máquina. Como conjunto de teste utilizou-se os dados desconhecidos e no final deste conjunto foram adicionadas 9 amostras que assinalam falhas manipuladas na máquina.

O diagrama da implementação deste método encontra-se na fig. 54.

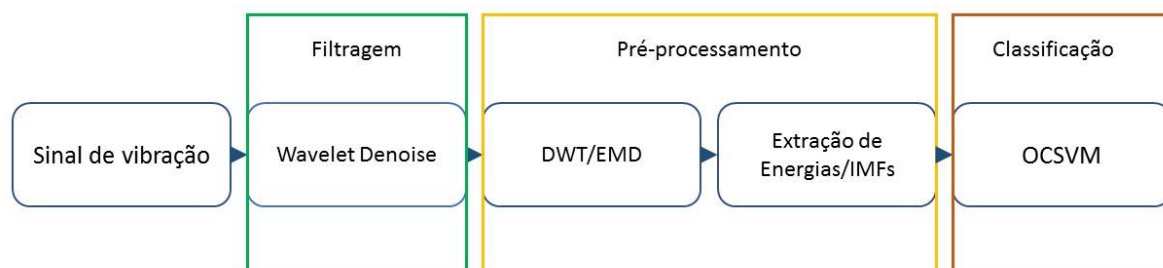


Figura 54-Diagrama de implementação DWT + FCM

As funções utilizadas para extração de energia utilizando a DWT e a EMD estão na fig. 55. A função EMD utilizada foi desenvolvida por Rilling and Flandrin ("Empirical Mode Decomposition," n.d.).

```
[C,L] = wavedec(New_X,level_dec,wavelet);
IMF = emd (New_X,OPTIONS);
[E_apr,E_det] = wenergy (C,L);
Energy = horzcat(E_apr,E_det);
```

*Figura 55-Funções utilizadas para obter as energias do sinal utilizando a DWT e a EMD*

Depois introduziram-se estas energias no classificador OCSVM. A função de treino e teste estão representadas na fig. 56. Na função de treino o número introduzido no quadrado a verde, indica o tipo de SVM, neste caso o número 2 indica que o modelo de treino gerado ("model") vai ser do tipo OCSVM.

```
model = svmtrain(label,matrix,[' -s 2 -t 2 -n ',num2str(best_C),' -g ', num2str(best_gamma)]);
[p_labels,acc,prob] = svmpredict(label_t,matrix_t, model);
```

*Figura 56-Funções de treino e teste da OCSVM*

A classificação produz valores de 1 e -1, em que 1 indica bom estado da máquina e -1 indica o mau estado.

Inicialmente, os dados de referência foram pré-processados pelo método DWT e as suas características introduzidas na OCSVM para classificação. Os resultados encontram-se nas figuras 57 e 58.

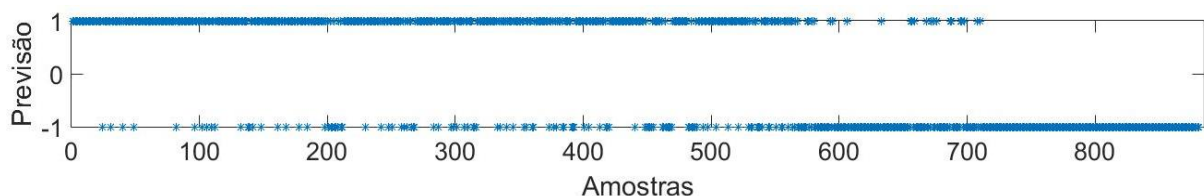


*Figura 57-Classificação OCSVM utilizando 8 energias extraídas pela DWT*

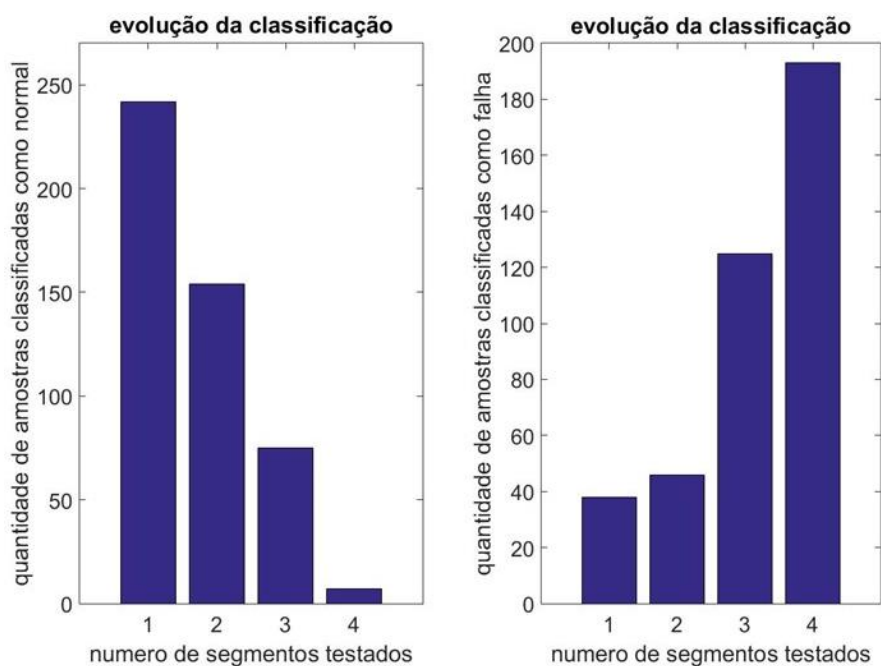


Figura 58-Evolução da classe normal e de falha na classificação da fig. 50

Para facilitar a interpretação das figuras que representam as classificações do algoritmo testado, retirou-se a quantidade de amostras nas classes normal (etiqueta 1) e falha (etiqueta -1), ao longo do conjunto de dados. Observando os gráficos das fig. 58, 60 e 62, a quantidade de amostras em boas condições diminui, verificando-se um aumento da classe falha, tal como era expectável.

Sabendo que nos dados de referência, ocorre uma falha no final da experiência teste-falha chega-se à conclusão que ocorreu uma classificação razoável. Observa-se um aumento gradual da classe -1, que indica a condição de falha. A classificação vai de encontro à informação que se possui a cerca das amostras utilizadas.

Utilizando o método de EMD, obtiveram-se resultados igualmente satisfatórios, como se pode ver nas fig. 59 e 61.

Observando as duas figuras consegue-se ver o aumento gradual da classe -1, a concentração de dados na classe 1 no início, e a concentração de dados na classe -1 no final da figura, como é de esperar levando em conta a informação acerca da falha nos dados de referência.

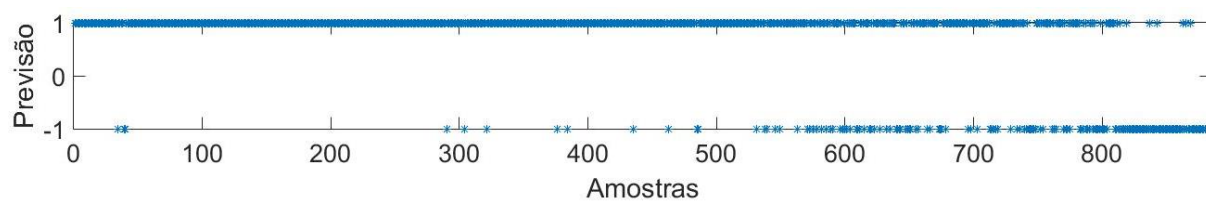


Figura 59-Classificação OCSVM com energias das 8 IMFs

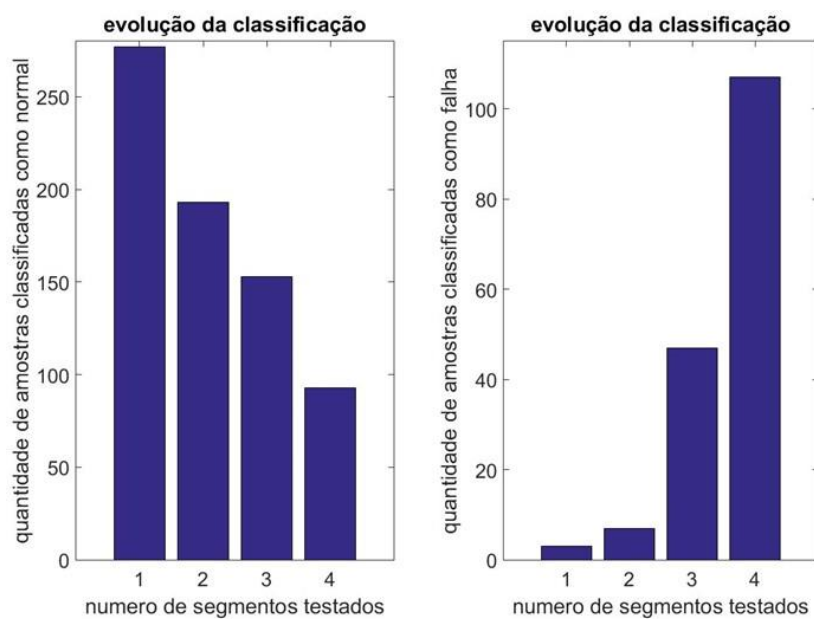


Figura 60-Evolução da classe normal e de falha na classificação da fig. 53

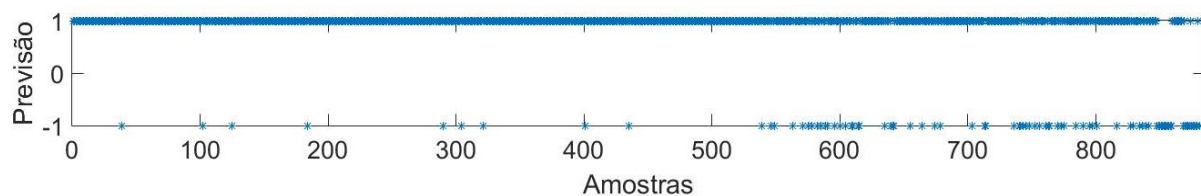


Figura 61-Classificação com energias das 2 primeiras IMFs

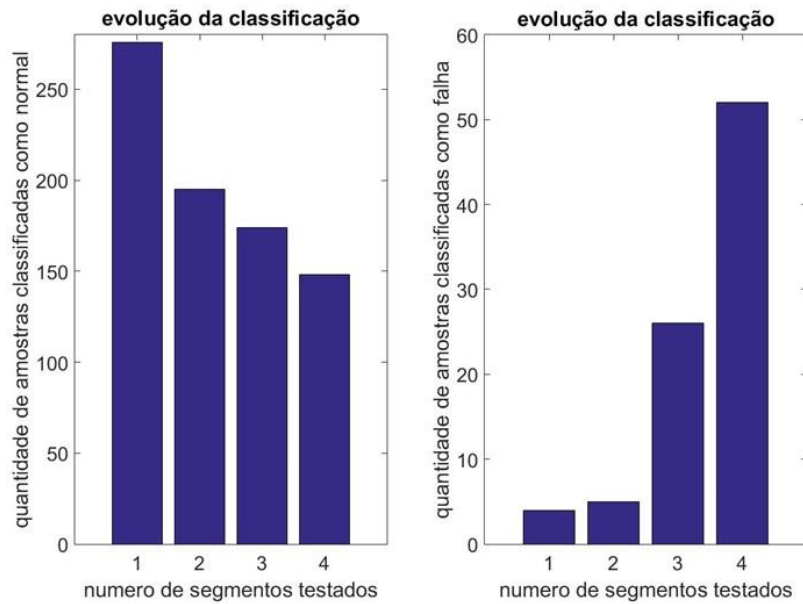


Figura 62-Evolução da classe normal e de falha na classificação da fig. 50

Observando a classificação das fig. 59 e 61, verifica-se que utilizando 8 IMFs no método EMD consegue-se realizar uma avaliação mais fiável do comportamento da máquina, comparativamente à utilização de apenas as duas primeiras IMFs.

Posto isto, foram realizados testes com os dados cedidos pela empresa Controlar.

Nas fig. 63 até à fig. 58 estão representados os testes realizados para os dois tipos de pré-processamento.

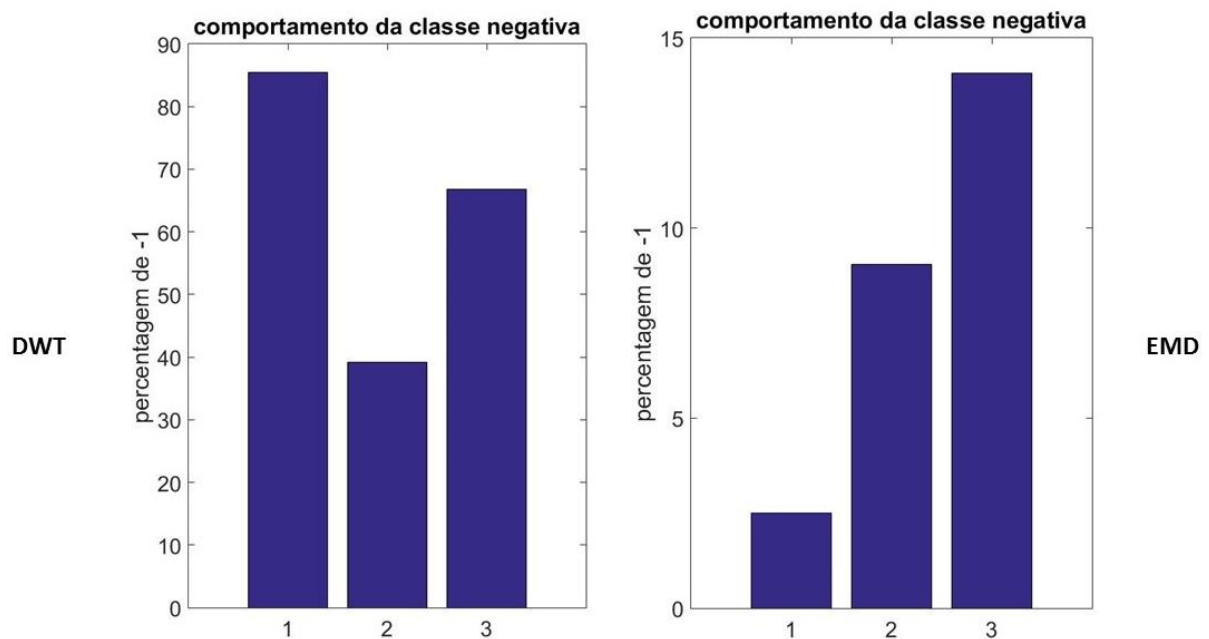


Figura 63-Evolução da classe -1, em cada eixo, para cada método de pré-processamento

Na fig. 63 encontra-se representada a evolução da classe -1 em percentagem, correspondente aos eixos X, Y e Z, respetivamente. A evolução da classe -1 representa a quantidade de amostras que foram classificadas como falha pela OCSVM. As fig. 64 e 65 dizem respeito à classificação ocorrida, em cada eixo individualmente, para cada um dos métodos.

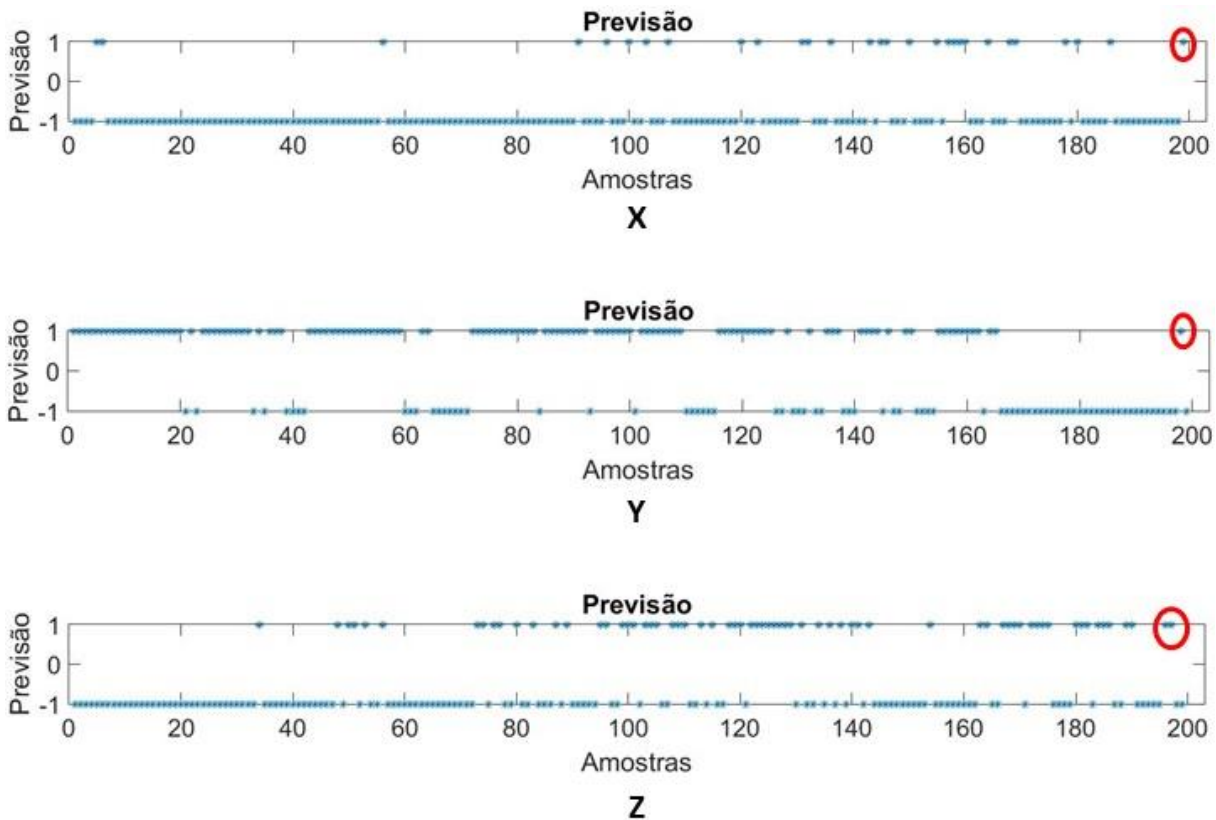


*Figura 64-Classificação dos dados da Controlar utilizando a DWT*

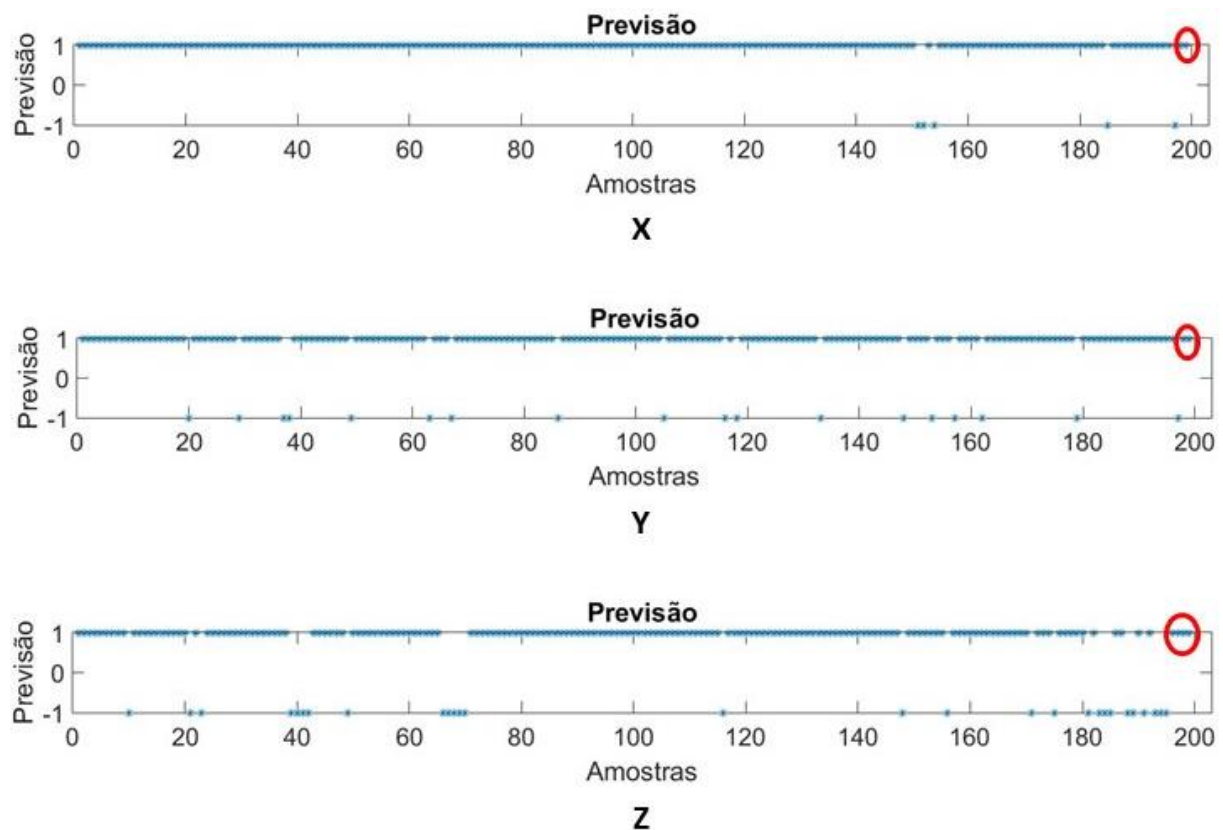


*Figura 65-Classificação dos dados da Controlar utilizando a EMD*

Chega-se à conclusão que ocorreu uma má classificação, pelo facto de que, em cada eixo, alguns dos últimos 9 elementos (que são falhas manipuladas da máquina) encontram-se na classe +1 (assinalado a vermelho), que é a classe que indica o bom estado da máquina.

A diferença de comportamento da classe -1 nos 3 eixos comprova a diferente informação que se pode adquirir de diferentes posições do acelerómetro, aquando da recolha do sinal de vibração da máquina.

#### 6.2.2 DWT/EMD + SVDD

Após os maus resultados obtidos na classificação dos dados da Controlar e, visto que o método SVDD é um método mais eficiente na classificação de uma só classe, pois fornece uma descrição melhor dos dados, optou-se por fazer esta a mesma implementação, mas utilizando o classificador SVDD com objetivo de conseguir melhores resultados.

O diagrama geral da implementação está na fig. 66.

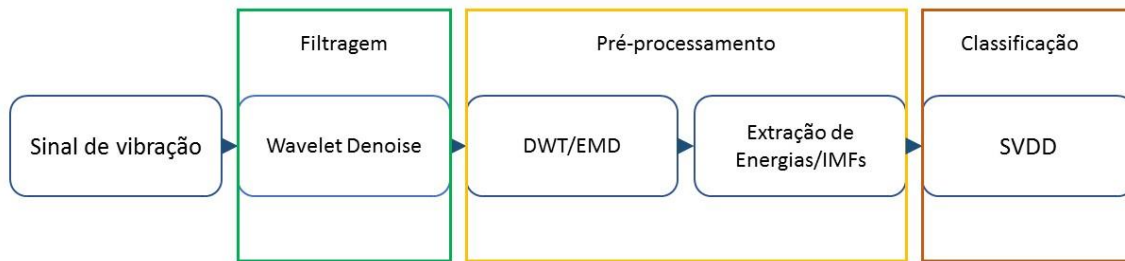


*Figura 66-Diagrama de implementação DWT/EMD + SVDD*

Para validar este método procederam-se a dois testes diferentes dos realizados até ao momento.

Em vez de se repartir os dados de referência em conjuntos de treino e teste com um número fixo de elementos, optou-se por fazer vários testes para se observar a evolução dos resultados da classificação. Começou-se com uma distribuição de 500 amostras para treino, e o restante para teste, e foram realizados 16 testes. Em cada teste acrescentaram-se 100 amostras ao conjunto de treino e, conseqüentemente, eram retiradas 100 ao conjunto de teste. Este procedimento encontra-se visível no fluxograma da fig. 67.

Como o conjunto de dados disponibilizados pela empresa é diferente e não contém o mesmo número de dados, o teste efetuado com esses dados é ligeiramente diferente, contudo tem o mesmo propósito.

Na fig. 68 apresenta-se o fluxograma utilizado para realizar os testes com os dados cedidos. Nesta experiência há apenas um conjunto de teste, no qual a condição (de bom estado ou falha) das amostras é desconhecida, apenas se sabe que no final deste conjunto, 9 amostras representam falhas. Há 6 conjuntos de treino, em que, o primeiro conjunto tem 17 amostras e, a cada teste, vão sendo adicionadas 10 amostras.

Para se ter uma ideia mais clara da classificação dos dados cedidos pelo algoritmo SVDD, foi realizado outro teste. Neste novo teste, apenas os sinais conhecidos (com e sem falha) foram utilizados. Assim, como treino foram utilizadas 50 amostras sem falha, e para teste formou-se um conjunto de 17 amostras sem falha mais a totalidade de amostras com falha (9 amostras). Depois foi contabilizado o número de falsos positivos e falsos negativos.

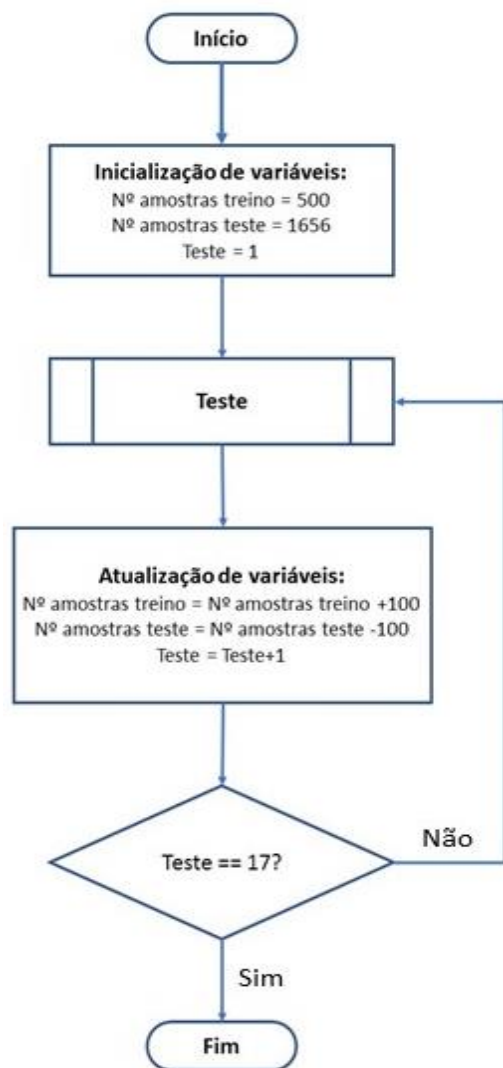


Figura 68- Procedimento de teste do conjunto de dados de referência

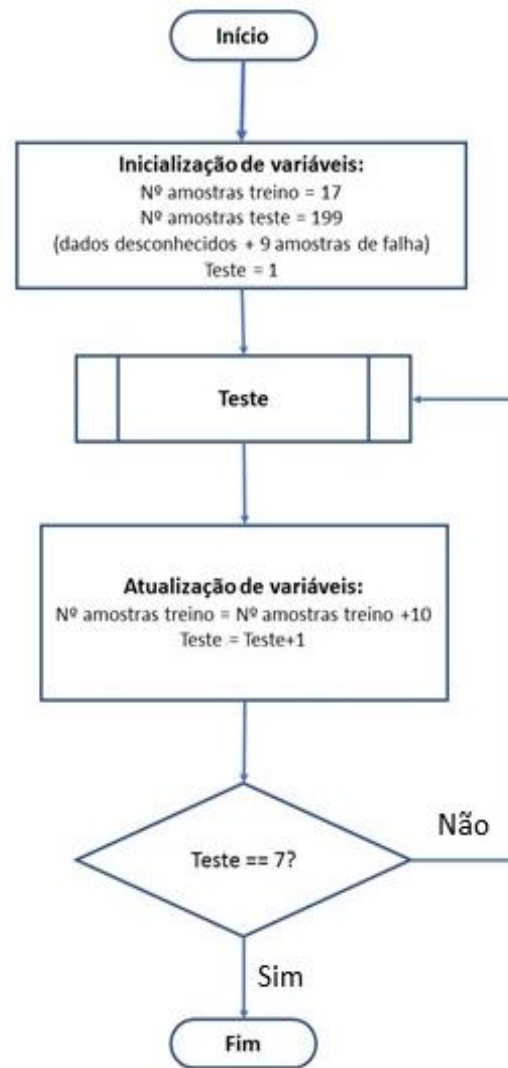


Figura 67-Procedimento de teste das amostras da Controlar

Dos testes efetuados foi extraída a evolução da classificação, a evolução da classe -1 e, para complementar a interpretação desta, a evolução da classe -1. É esperado que a classe -1 aumente com a diminuição do conjunto de teste pois, como a falha se encontra nas amostras finais, com a diminuição dos elementos no conjunto de teste a falha vai ficar mais evidente.

A função de treino e teste da SVDD estão representadas na fig. 69. Recordando o que foi dito em 6.2.1, o número introduzido no quadrado a verde, indica o tipo de SVM, assim o número 5 assinala que o modelo de treino gerado ("model") é do tipo SVDD.

```

model = svmtrain(label,matrix,[' -s 5 -t 2 -n ',num2str(best_C),' -g ', num2str(best_gamma)]);
[p_labels,acc,prob] = svmpredict(label_t,matrix_t, model);
  
```

Figura 69-Funções de treino e teste da OCSVM

Tendo presente os testes realizados com este algoritmo para os dados de referência (fluxograma da fig. 67) os resultados das classificações estão nas figuras seguintes.

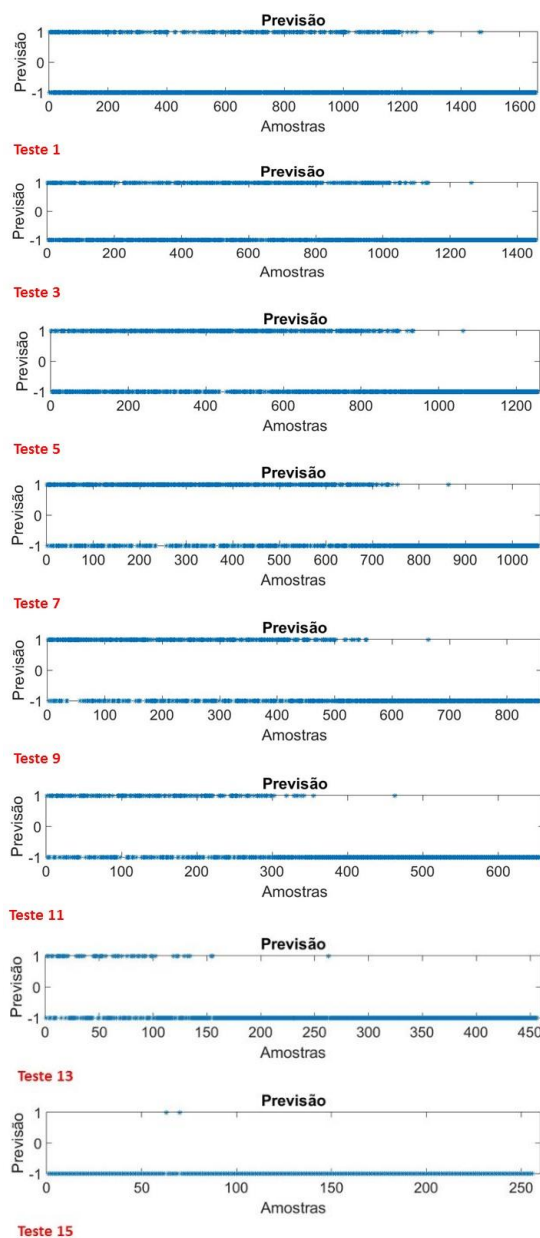


Figura 70-Testes utilizando a DWT com a wavelet mãe db8

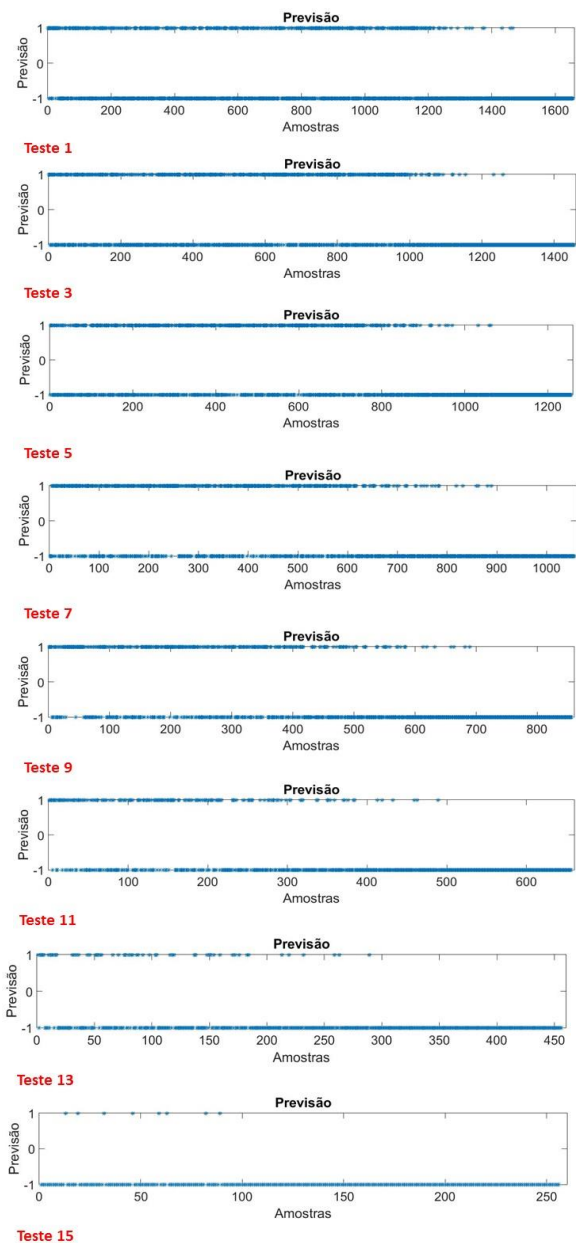


Figura 71-Testes utilizando a DWT com a wavelet mãe haar

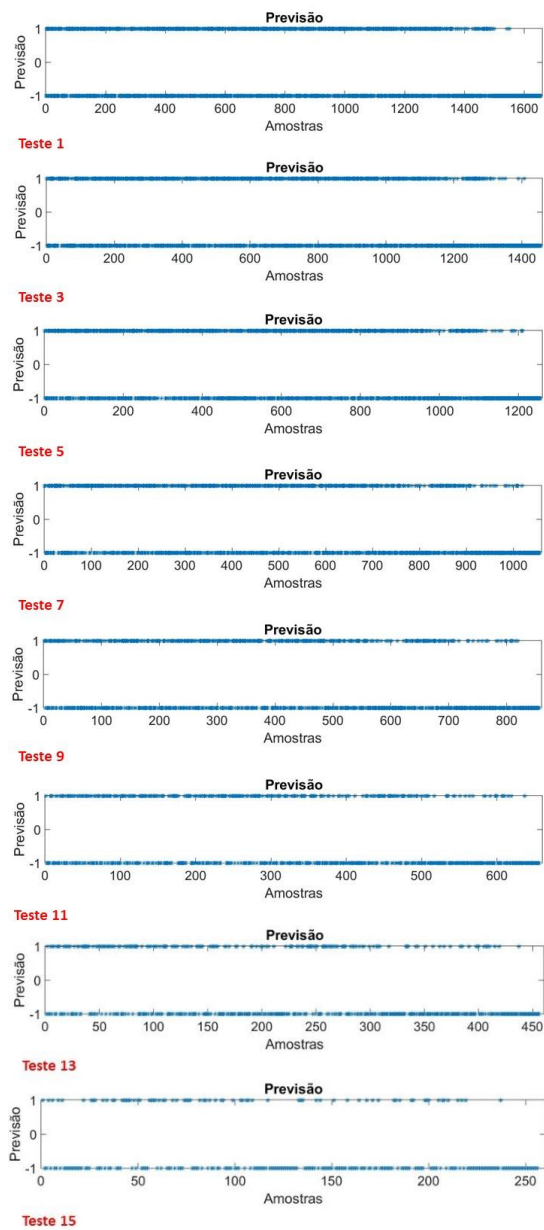


Figura 72-Testes utilizando a EMD com 8 energias

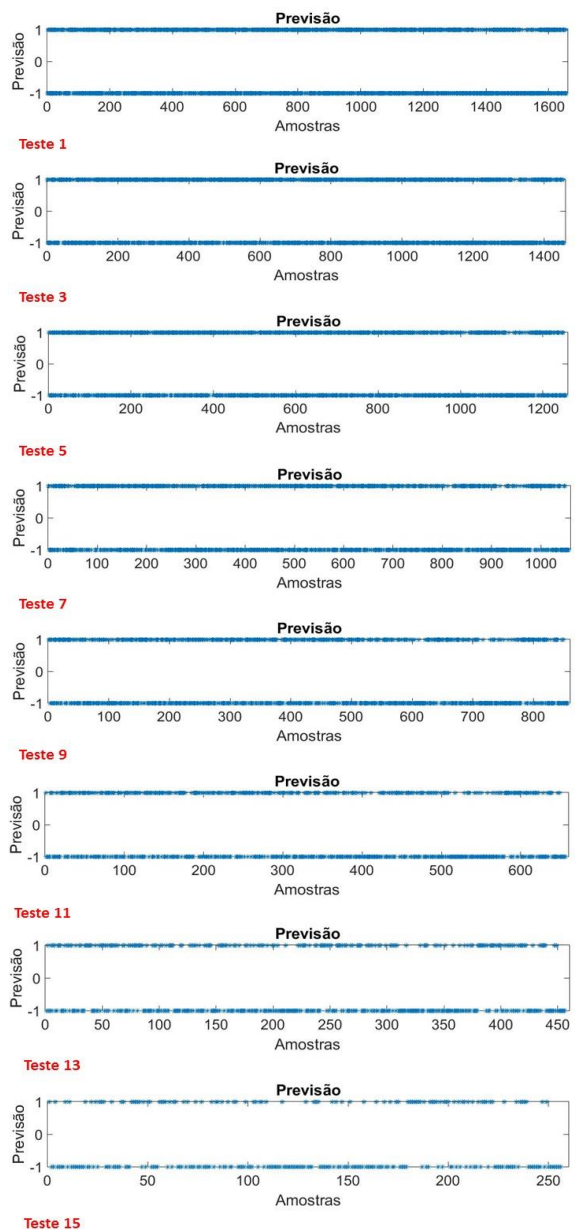


Figura 73-Testes utilizando a EMD com 2 energias

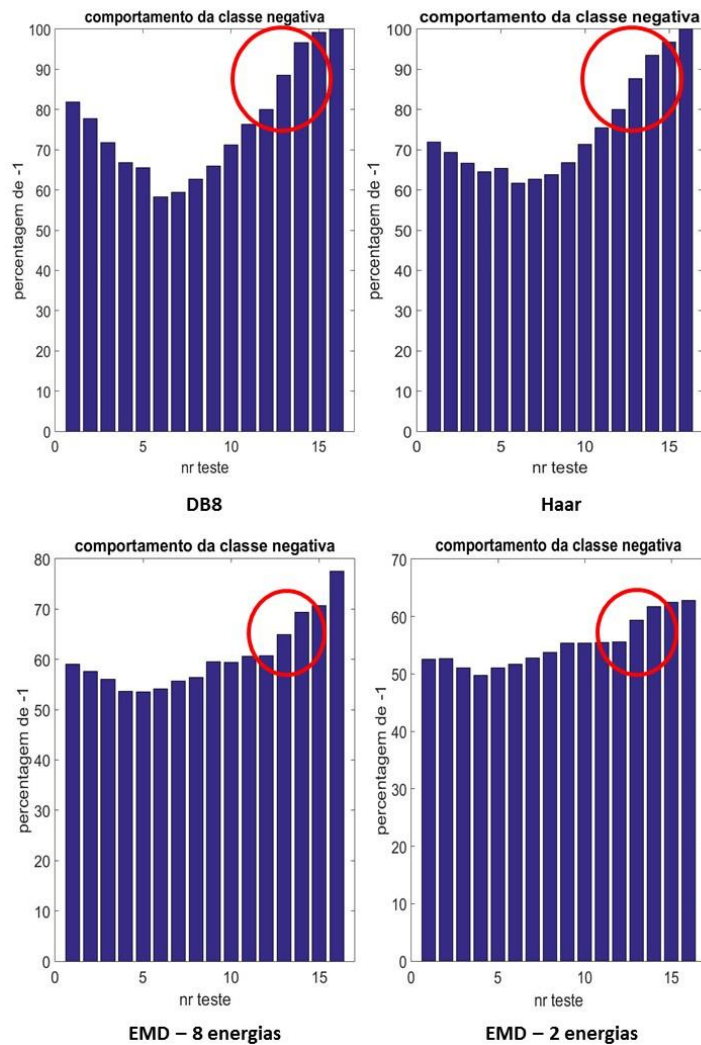


*Figura 74-Evolução da classe -1 ao longo dos 16 testes para cada método de pré-processamento*

Nas figuras 70 a 73 estão representados 8 dos 16 testes feitos, para as 2 técnicas de processamento de sinal referidas. Estas figuras mostram a evolução da classificação feita pela SVDD. Esta classificação é mais perceptível na fig. 74 onde se pode ver a evolução da classe -1, classe que representa a condição de falha no rolamento, para cada técnica utilizada.

Observando os gráficos da fig. 74 conclui-se que, para todos os métodos, o comportamento da classe -1 é idêntico. No início é-lhe atribuída uma grande taxa, que vai diminuindo com o tempo. Isto é devido ao facto de a SVDD ter um comportamento adaptativo. Com mais dados de treino consegue obter uma melhor classificação. No entanto nos testes finais, com a diminuição dos dados de teste, a falha torna-se mais evidente e, consequentemente, a taxa da classe -1 aumenta, conforme esperado.

Os resultados obtidos são satisfatórios, do ponto de vista teórico, visto que a classificação ocorreu conforme a informação que se tem acerca dos dados, porém ao analisar as fig. 70 a 73 verifica-se que em cada teste existem muitos falsos negativos. Isto significa que, num ambiente industrial real, a máquina iria assinalar falha numa situação de condição de funcionamento normal, o que levaria a falsos alertas e que originaria o descrédito do sistema de previsão/prevenção de avaria.

Os gráficos da fig. 74 retratam bem a tendência dos dados e do crescimento da classe -1. Também se pode ver um aumento significativo entre os testes 11 a 14, assinalado a vermelho, por essa razão, quando a falha começa a manifestar-se, mais claramente, algures entre a amostra 1600 e 1800.

Comparando os métodos de pré-processamento utilizados, a EMD e a DWT, verificou-se que a primeira demora significativamente mais tempo que a segunda para efetuar o pré-processamento do sinal.

Na utilização da EMD com 2 energias conseguimos uma avaliação razoável com menos custo computacional, quando comparado com a utilização de 8 energias, contudo não é tão precisa e confiável.

Considerando os resultados obtidos, procedeu-se ao teste da SVDD com os dados cedidos pela empresa Controlar. O procedimento do teste encontra-se no fluxograma da fig. 68 e os resultados obtidos encontram-se na fig. 75.

Analisando as três imagens, conseguem-se ver semelhanças na evolução da classe -1, nos eixos do Z e dos X, para as 3 condições de pré-processamento utilizadas.

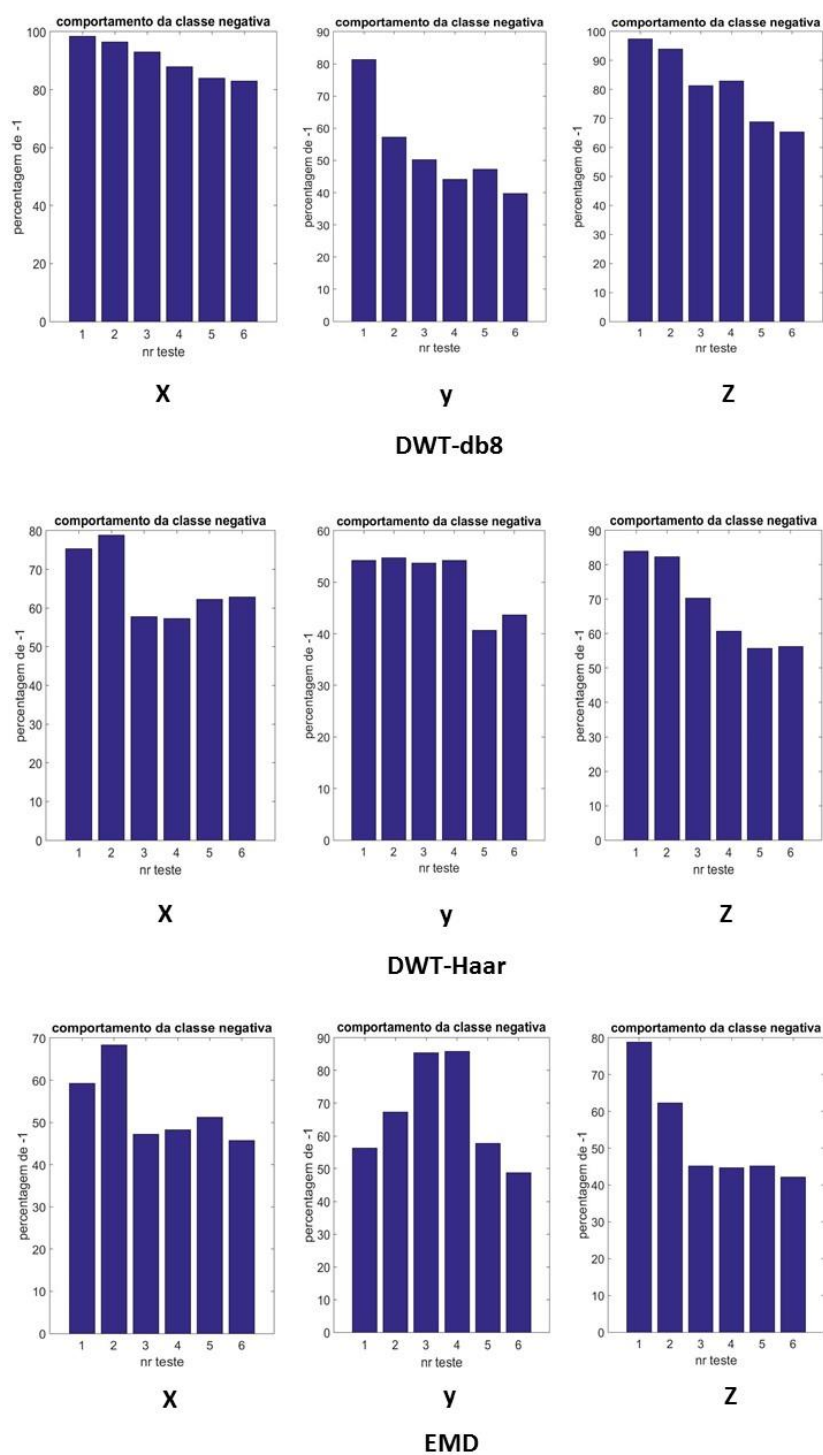


Figura 75-Evolução da classe -1 em cada eixo para cada método de pré-processamento

Como foi referido, o número de falsos positivos e falsos negativos foi contabilizado para verificar a qualidade da classificação. Esta contagem encontra-se nas tabelas 5, 6 e 7.

Tabela 7-Classificação da SVDD em cada eixo, utilizando a DWT (db8)

|   |               |        |       |
|---|---------------|--------|-------|
| X | Classificação | Classe |       |
|   |               | normal | falha |
|   |               | normal | 2     |
| Y | Classificação | Classe |       |
|   |               | normal | falha |
|   |               | normal | 2     |
| Z | Classificação | Classe |       |
|   |               | normal | falha |
|   |               | normal | 4     |

Tabela 7-Classificação da SVDD em cada eixo, utilizando a DWT (haar)

|   |               |        |       |
|---|---------------|--------|-------|
| X | Classificação | Classe |       |
|   |               | normal | falha |
|   |               | normal | 6     |
| Y | Classificação | Classe |       |
|   |               | normal | falha |
|   |               | normal | 2     |
| Z | Classificação | Classe |       |
|   |               | normal | falha |
|   |               | normal | 2     |

Tabela 7-Classificação da SVDD em cada eixo, utilizando a EMD

|   |               |        |       |
|---|---------------|--------|-------|
| X | Classificação | Classe |       |
|   |               | normal | falha |
|   |               | normal | 8     |
| Y | Classificação | Classe |       |
|   |               | normal | falha |
|   |               | normal | 0     |
| Z | Classificação | Classe |       |
|   |               | normal | falha |
|   |               | normal | 7     |

Analisando estas tabelas verifica-se que a classificação não é feita corretamente. Há uma quantidade grande de falsos negativos em todos os testes realizados, concluindo-se que a classificação não é fiável. Por isso, a utilização deste sistema num ambiente industrial para monitorizar máquinas industriais não é aconselhável, visto que, o sistema vai dar alerta de falhas inexistentes.

### 6.3 Conclusão

Devido aos resultados insatisfatórios obtidos nos capítulos anteriores, neste capítulo implementaram-se dois novos métodos.

No pré-processamento, além da utilização da DWT, foi introduzido uma nova técnica, o EMD, que decompõem o sinal em monocomponentes (IMFs). Estes dois algoritmos foram utilizados para retirar 8 energias, no caso da DWT, e 8/2 IMFs, no caso do EMD, em cada amostra de vibração.

Para realizar a classificação foram implementados dois classificadores adaptativos e não supervisionados, o OCSVM e SVDD.

Nas experiências concretizadas com o classificador OCSVM utilizando os dados de referência, a classificação obtida foi razoável, conseguindo-se observar o aumento gradual da classe de falha, comprovando a informação obtida desse conjunto de dados.

Os testes realizados com os dados cedidos pela empresa utilizando a OCSVM permitiram observar a diferente informação que se pode extrair dos vários eixos. No entanto, os resultados obtidos não foram satisfatórios pois houve má classificação. Assim, foi implementado o classificador SVDD por ser mais eficiente na classificação de uma só classe.

Com a realização de vários testes utilizando a SVDD, com os dados de referência observa-se a evolução da falha e confirma-se o comportamento adaptativo do algoritmo.

Teoricamente os resultados são satisfatórios pois a classificação ocorre como esperado, tendo em conta a experiência teste-falha que originou os dados analisados. No entanto, utilizando este método num meio industrial não seria adequado pois, a classificação contém muitos falsos negativos, levando a falsos alertas e manutenções desnecessárias, descreditando o método implementado.

Introduzindo os dados fornecidos pela empresa na SVDD, as classificações continham uma grande quantidade de falsos negativos em todos os testes, concluindo-se que a abordagem não é fiável para ser utilizada num ambiente industrial.

## 7. CONCLUSÕES E TRABALHO FUTURO

Neste capítulo são apresentadas as conclusões finais deduzidas a partir da avaliação de todo o trabalho desenvolvido.

Para além da conclusão, também são referidos alguns pontos, como trabalho futuro, para dar seguimento ao trabalho realizado.

### 7.1 Conclusão

Hoje em dia, a indústria está a passar por um grande processo de evolução e adaptação, devido à 4ª revolução industrial, marcada pela utilização das tecnologias de informação nos processos industriais que permitem a análise e interpretação de grandes conjuntos de dados gerados por diversas fontes.

Numa empresa industrial a manutenção é um processo essencial para o seu bom funcionamento. Este pode ser melhorado, transformando-a num processo autónomo, eficiente e confiável através da união de uma abordagem científica com as tecnologias de informação.

Mediante a monitorização de vários parâmetros de uma máquina industrial é possível desenvolver um diagnóstico do estado da máquina e detetar falhas prematuramente, prevenindo a sua avaria.

A presente dissertação faz parte de uma linha de investigação da empresa Controlar que tem como finalidade desenvolver uma solução para a monitorização de máquinas rotativas, aplicando técnicas de *Data Mining* e *Machine Learning* a dados de vibração.

O desenvolvimento foi executado iterativamente, com base em artigos e na pesquisa realizada, procurando sempre satisfazer o objetivo principal.

Ao longo do desenvolvimento desta dissertação foram realizadas várias experiências recorrendo a vários algoritmos das áreas de filtragem, processamento de sinal, data mining e machine learning.

Todos os testes e algoritmos desenvolvidos encontram-se na tabela 8.

*Tabela 8-Experiências implementados*

| Implementação | Filtragem | Normalização                                                               | Pré-processamento                                                                      | Classificador | Conjunto de dados                                                                          |
|---------------|-----------|----------------------------------------------------------------------------|----------------------------------------------------------------------------------------|---------------|--------------------------------------------------------------------------------------------|
| 1             | BSS       | <ul style="list-style-type: none"> <li>Standard</li> </ul>                 | <ul style="list-style-type: none"> <li>Curtograma rápido e análise envelope</li> </ul> | —             | <ul style="list-style-type: none"> <li>Referência</li> </ul>                               |
| 2             | WD        | <ul style="list-style-type: none"> <li>Standard</li> </ul>                 | <ul style="list-style-type: none"> <li>WTP (db8)</li> </ul>                            | —             | <ul style="list-style-type: none"> <li>Referência</li> <li>Falhas específicas</li> </ul>   |
| 3             | WD        | <ul style="list-style-type: none"> <li>Standard</li> </ul>                 | <ul style="list-style-type: none"> <li>DWT (db8)</li> </ul>                            | SVM           | <ul style="list-style-type: none"> <li>Referência</li> <li>Cedidos pela empresa</li> </ul> |
| 4             | WD        | <ul style="list-style-type: none"> <li>Standard</li> </ul>                 | <ul style="list-style-type: none"> <li>DWT (db8)</li> </ul>                            | FCM           | <ul style="list-style-type: none"> <li>Referência</li> </ul>                               |
| 5             | WD        | <ul style="list-style-type: none"> <li>Artigo (M. et al., 2012)</li> </ul> | <ul style="list-style-type: none"> <li>DWT (db8 e haar)</li> <li>EMD</li> </ul>        | OCSVM         | <ul style="list-style-type: none"> <li>Referência</li> <li>Cedidos pela empresa</li> </ul> |
| 6             | WD        | <ul style="list-style-type: none"> <li>Artigo (M. et al., 2012)</li> </ul> | <ul style="list-style-type: none"> <li>DWT (db8 e haar)</li> <li>EMD</li> </ul>        | SVDD          | <ul style="list-style-type: none"> <li>Referência</li> <li>Cedidos pela empresa</li> </ul> |

A extração da energia dos dados de vibração, aliado ao classificador SVDD foi o melhor método empregue, tanto para os dados de referência como para dos dados fornecidos pela empresa de acolhimento, atingindo os melhores resultados. Além disso, a SVDD é o classificador mais adequado para realizar a monitorização das máquinas industriais num ambiente real, por ser um classificador de uma classe, em que normalmente os únicos dados disponíveis de uma máquina industrial são dados que refletem o bom funcionamento da mesma e permite uma descrição melhor dos dados no espaço de características.

Os resultados obtidos em cada implementação apresentada na tabela 8 encontra-se na tabela 9.

Tabela 9-Resultados obtidos em cada experiência realizada

| Implementação | Algoritmos                                 | Conjunto de dados                    | Resultado |
|---------------|--------------------------------------------|--------------------------------------|-----------|
| 1             | BSS + Curtograma rápido e análise envelope | • Referência                         | —         |
| 2             | WD + WPT(db8)                              | • Referência<br>• Falhas específicas | Razoável  |
| 3             | WD + DWT (db8) + SVM                       | • Referência                         | Razoável  |
|               |                                            | • Cedidos pela empresa               | Medíocre  |
| 4             | WD + DWT (db8) + FCM                       | • Referência                         | Medíocre  |
| 5             | WD + DWT (db8, haar)/EMD + OCSVM           | • Referência                         | Razoável  |
|               |                                            | • Cedidos pela empresa               | Medíocre  |
| 6             | WD + DWT (db8, haar)/EMD + SVDD            | • Referência                         | Razoável  |
|               |                                            | • Cedidos pela empresa               | Medíocre  |

Um dos problemas que surgiu nesta dissertação foi o problema dos falsos positivos e negativos que pode ser resolvido, introduzindo no classificador, um conjunto de treino mais representativo da máquina a monitorizar, onde seja levado em consideração o meio e a sequência de aquisição dos dados.

Os maus resultados gerais aquando da utilização dos dados cedidos pela Controlar devem-se ao facto de os dados serem manipulados, não controlados, não documentados e onde não se tem conhecimento da sua sequência de aquisição.

Por fim, o objetivo principal desta dissertação foi atingido pois foram aplicadas diversas metodologias no pré-processamento e análise dos dados de vibração, nomeadamente as técnicas de *data mining* e *machine learning*, e foi avaliado o seu desempenho recorrendo a dados de referência e a dados de uma máquina industrial real.

## 7.2 Trabalho futuro

O principal problema do sistema desenvolvido é a presença de falsos negativos na classificação. A melhoria da fiabilidade do sistema passa pelo desenvolvimento de um método

capaz de distinguir os falsos negativos, para que não ocorram falsos alarmes que levem à manutenção desnecessária da máquina.

Para complementar o processo de manutenção preventiva seria interessante implementar métodos que realizem o prognóstico da máquina, como por exemplo o algoritmo SVR. Assim, através do cálculo do tempo de vida dos componentes monitorizados é possível fazer uma estimativa/previsão do período em que vai ocorrer a sua falha.

Adicionalmente, para facilitar a consulta de toda a informação pelas entidades da empresa, seria também interessante o desenvolvimento de uma base de dados que guardasse toda a informação gerada pelo sistema, e emita alertas quando as máquinas apresentem indícios prematuros de falha. Assim, será possível o planeamento adequado da sua manutenção.

## BIBLIOGRAFIA

- Aherwar, A. (2013). An investigation on gearbox fault detection using vibration analysis techniques: a review. *Australian Journal of Mechanical Engineering*, 10(2), 169. <https://doi.org/10.7158/M11-830.2012.10.2>
- Ahmad, R., & Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers and Industrial Engineering*, 63(1), 135–149. <https://doi.org/10.1016/j.cie.2012.02.002>
- Amaral, L. M. (2016). O conceito de reindustrialização, Indústria 4.0 e a política industrial para o século XXI - O caso português, 1–40. Retrieved from [http://cip.org.pt/wp-content/uploads/2016/06/2016-05-17\\_LMAMARAL\\_Leiria.pdf](http://cip.org.pt/wp-content/uploads/2016/06/2016-05-17_LMAMARAL_Leiria.pdf)
- Amir, R. Bin, Gul, S. T., & Khan, A. Q. (2016). A Comparative Analysis of Classical and One Class SVM Classifiers for Machine Fault Detection Using Vibration Signals.
- Bagheri, B., Ahmadi, H., & Labbafi, R. (2010). Application of Data Mining and Feature Extraction on Intelligent Fault Diagnosis by Artificial Neural Network and k-Nearest Neighbor.
- Baqqar, M., Ahmed, M., & Gu, F. (2011). Data mining for gearbox condition monitoring. *The 17th International Conference on Automation and Computing*, (September), 138–142.
- Benkedjouh, T., Medjaher, K., Zerhouni, N., & Rechak, S. (2012). Fault prognostic of bearings by using support vector data description. *Prognostics and Health Management (PHM), 2012 IEEE Conference on*, 1–7. <https://doi.org/10.1109/ICPHM.2012.6299511>
- Cao, S.-Q., Ma, X., Zhang, Y., Luo, L., & Yi, F. (2015). A Fault Diagnosis Method Based on Semi - Supervised Fuzzy C - Means Cluster Analysis. *International Journal on Cybernetics & Informatics*, 4(2), 281–289. <https://doi.org/10.5121/ijci.2015.4227>
- Chang, C.-C., & Lin, C.-J. (2015). LIBSVM FAQ. Retrieved February 4, 2017, from <http://www.csie.ntu.edu.tw/~cjlin/libsvm/faq.html>
- Chang, C.-C., & Lin, C.-J. (2016). LIBSVM -- A Library for Support Vector Machines. Retrieved January 1, 2017, from <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- Chang, C., & Lin, C. (2011). A Library for Support Vector Machines. *ACM Transactions on Interlligent Systems and Technology (TIST)*, 2(3), 39. <https://doi.org/10.1145/1961189.1961199>
- Chen, M., Mao, S., & Liu, Y. (2014). Big Data : A Survey, 171–209. <https://doi.org/10.1007/s11036-013-0489-0>
- Chorowski, J., Wang, J., & Zurada, J. M. (2014). Review and performance comparison of SVM- and ELM-based classifiers. *Neurocomputing*, 128, 507–516. <https://doi.org/10.1016/j.neucom.2013.08.009>
- Courrech, J., & Guudet, M. (1996). Envelope analysis - the key to rolling-element bearing diagnosis. *Brüel & Kjaer*, 8.
- Dan, Z., & Wentao, S. (2014). 48 . MED and WPT based technique for bearings fault detection, 103–110.
- Deák, K., Kocsis, I., Vámosi, A., & Keviczki, Z. (2014). FAILURE DIAGNOSTICS WITH SVM IN MACHINE MAINTENANCE ENGINEERING, (May).
- Delft, T. U., & Magnificus, R. (2001). One-class classification.
- Dong, S., Yin, S., Tang, B., Chen, L., & Luo, T. (2014). Bearing degradation process prediction based on the support vector machine and Markov model. *Shock and Vibration*, 2014. <https://doi.org/10.1155/2014/717465>

- Dr. Saed Sayad. (2017). An Introduction to Data Mining. Retrieved January 1, 2017, from [http://www.saedsayad.com/k\\_nearest\\_neighbors\\_reg.htm](http://www.saedsayad.com/k_nearest_neighbors_reg.htm)
- Empirical Mode Decomposition. (n.d.). Retrieved from <http://perso.ens-lyon.fr/patrick.flandrin/emd.html>
- Fang, X., & Luo, H. (2009). An Improved Hilbert Transform for Nonlinear Vibration, (May).
- Fangfang, X., Long, Y., Xiucai, Z., & Wenan, H. (2015). Intelligent fault diagnosis based on support vector machine, 192–196.
- Fei, S. wei. (2017). Fault Diagnosis of Bearing Based on Wavelet Packet Transform-Phase Space Reconstruction-Singular Value Decomposition and SVM Classifier. *Arabian Journal for Science and Engineering*, 42(5), 1967–1975. <https://doi.org/10.1007/s13369-016-2406-x>
- George Dallas. (2014). wavelets @ georgemdallas.wordpress.com. Retrieved January 1, 2017, from <https://georgemdallas.wordpress.com/2014/05/14/wavelets-4-dummies-signal-processing-fourier-transforms-and-heisenberg/>
- Gokhale, M. Y., & Khanduja, D. K. (2010). Time Domain Signal Analysis Using Wavelet Packet Decomposition Approach. *International Journal of Communications, Network and System Sciences*, 3(3), 321–329. <https://doi.org/10.4236/ijcns.2010.33041>
- Goyal, D., & Pabla, B. S. (2015). The Vibration Monitoring Methods and Signal Processing Techniques for Structural Health Monitoring: A Review. *Archives of Computational Methods in Engineering*. <https://doi.org/10.1007/s11831-015-9145-0>
- Graps, A. (2004). Introduction to the Discrete Wavelet Transform (DWT), 3, 1–8.
- Grosse, C. U., & Reinhardt, H. W. (2002). Signal conditioning in acoustic emission analysis using wavelets. *The E-Journal of Nondestructive Testing*, 7(9), 1–8. Retrieved from <http://www.ndt.net/article/v07n09/08/08.htm>
- Gutierrez-Osuna, R. (2002). Short-time Fourier analysis and synthesis. “*Introduction to Speech Processing*”, Texas A&M, 1–26. Retrieved from <http://research.cs.tamu.edu/prism/lectures/sp/l6.pdf>
- Haar, D. (1998). 3. Analyse de, 26–48.
- Han, J., Rafii, Z., & Pardo, B. (2017). Audio Source Separation. Retrieved January 1, 2017, from <http://music.cs.northwestern.edu/research.php>
- Hua, J. H. J. (2008). Study on the Application of Rough Sets Theory in Machine Learning. *2008 Second International Symposium on Intelligent Information Technology Application*, 1, 192–196. <https://doi.org/10.1109/IITA.2008.154>
- Huang, H., Ouyang, H., Gao, H., Guo, L., Li, D., & Wen, J. (2016). A Feature Extraction Method for Vibration Signal of Bearing Incipient Degradation, (3), 149–159. <https://doi.org/10.1515/msr-2016-0018>
- Hwang, D.-H., Youn, Y.-W., Sun, J.-H., Choi, K.-H., Lee, J.-H., & Kim, Y.-H. (2015). Support Vector Machine Based Bearing Fault Diagnosis for Induction Motors Using Vibration Signals. *Journal of Electrical Engineering & Technology*, 10(4), 1558–1565. <https://doi.org/10.5370/jeet.2015.10.4.1558>
- Javed, K., Gouriveau, R., & Zerhouni, N. (2015). Enabling Health Monitoring Approach Based on Vibration Data for Accurate Prognostics, 62(1), 647–656.
- Jerome Antoni. (2015a). Fast Kurtogram. Retrieved January 1, 2017, from <https://www.mathworks.com/matlabcentral/fileexchange/48912-fast-kurtogram>
- Jerome Antoni. (2015b). Fast Kurtogram. Retrieved January 1, 2017, from <https://www.mathworks.com/matlabcentral/fileexchange/48912-fast-kurtogram?focused=3875917&tab=function>
- Jia, G., Yuan, S., & Tang, C. (2011). Fault Diagnosis of Roller Bearing Based on PCA and Multi-

- class Support Vector Machine, 198–205.
- Jing, Z., Hai, H., & Kaixing, H. (2014). Blind source separation of vibration signals for fault diagnosis of power transformers, 412–417.
- Joon, H., Eom, D., & Kim, S. (2005). One-class support vector machines — an application in machine fault detection and classification, 48, 395–408. <https://doi.org/10.1016/j.cie.2005.01.009>
- Kamaras, K., Ili, V. A. C., & Dimitrakopoulos, I. (2016). Bearings using Spectral Kurtosis and Envelope Analysis Vibration Analysis of Rolling Element, 1–14.
- Khan, M. A. S. K., Radwan, T. S., & Rahman, M. A. (2007). Real-time implementation of wavelet packet transform-based diagnosis and protection of three-phase induction motors. *IEEE Transactions on Energy Conversion*, 22(3), 647–655. <https://doi.org/10.1109/TEC.2006.882417>
- Konar, P., Sil, J., & Chattopadhyay, P. (2015). Knowledge extraction using data mining for multi-class fault diagnosis of induction motor. *Neurocomputing*, 166, 14–25. <https://doi.org/10.1016/j.neucom.2015.04.040>
- Konar, P., Sil, J., & Chattopadhyay, P. (2015). Neurocomputing Knowledge extraction using data mining for multi-class fault diagnosis of induction motor, 166, 14–25.
- Li, L., Qu, L., & Liao, X. (2007). Haar wavelet for machine fault diagnosis. *Mechanical Systems and Signal Processing*, 21(4), 1773–1786. <https://doi.org/10.1016/j.ymssp.2006.07.006>
- Liu, H., Li, L., & Ma, J. (2016). Rolling Bearing Fault Diagnosis Based on STFT-Deep Learning and Sound Signals. *Shock and Vibration*, 2016. <https://doi.org/10.1155/2016/6127479>
- Liu, H., Liu, C., & Huang, Y. (2011). Adaptive feature extraction using sparse coding for machinery fault diagnosis. *Mechanical Systems and Signal Processing*, 25(2), 558–574. <https://doi.org/10.1016/j.ymssp.2010.07.019>
- Lotfan, S., Salehpour, N., Adiban, H., & Mashroutechi, A. (2015). Bearing fault detection using fuzzy C-means and hybrid C-means-subtractive algorithms. *IEEE International Conference on Fuzzy Systems, 2015–Novem*. <https://doi.org/10.1109/FUZZ-IEEE.2015.7338049>
- Loutas, T., & Kostopoulos, V. (2012). Utilising the Wavelet Transform in Condition-Based Maintenance : A Review with Applications. *Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology*, 273–312.
- M., P., S., M., A., F., & L., G. (2012). External condition removal in bearing diagnostics through EMD and One-Class SVM.
- MathWorks. (2017). Wavelet Denoising. Retrieved January 1, 2017, from <https://www.mathworks.com/help/wavelet/ug/wavelet-denoising.html>
- Mobley, R. K. (2001). *Plant Engineer's Handbook*. Retrieved from [https://books.google.pt/books?id=9YkqH5HgSgwC&pg=SA40-PA117&lpg=SA40-PA117&dq=vibration+acquisition+sequence+for+monitoring+systems&source=bl&ots=UDKUfbpdp1&sig=165SOKzOlqbaEXdM5gxSwClhBZU&hl=pt-PT&sa=X&ved=0ahUKEwjyz\\_3zyu3UAhXJuRoKHRbIAOYQ6AEIUDAE#v=onepage&q=vibration+acquisition+sequence+for+monitoring+systems&f=false](https://books.google.pt/books?id=9YkqH5HgSgwC&pg=SA40-PA117&lpg=SA40-PA117&dq=vibration+acquisition+sequence+for+monitoring+systems&source=bl&ots=UDKUfbpdp1&sig=165SOKzOlqbaEXdM5gxSwClhBZU&hl=pt-PT&sa=X&ved=0ahUKEwjyz_3zyu3UAhXJuRoKHRbIAOYQ6AEIUDAE#v=onepage&q=vibration+acquisition+sequence+for+monitoring+systems&f=false)
- Moshe, D. (2004). Denoising using wavelets.
- Olshausen, B. A. (2000). Aliasing 1 Nyquist sampling theorem, 1–6.
- Pan, Y. N., Chen, J., & Dong, G. M. (2009). A hybrid model for bearing performance degradation assessment based on support vector data description and fuzzy c-means. *Proceedings of the Institution of Mechanical Engineers Part C-Journal of Mechanical Engineering Science*, 223(11), 2687–2695. <https://doi.org/10.1243/09544062jmes1447>
- Paulo, S. (2013). Aplicação da Transformada de Hilbert-Huang na Análise das Vibrações dos

Motores de Indução de Máquinas Ferramentas José Roberto Marques Aplicação da Transformada de Hilbert-Huang na Análise das Vibrações dos Motores de Indução de Máquinas Ferramentas São .

- Peng, Z. K., Tse, P. W., & Chu, F. L. (2005). A comparison study of improved Hilbert – Huang transform and wavelet transform : Application to fault diagnosis for rolling bearing, *19*, 974–988. <https://doi.org/10.1016/j.ymssp.2004.01.006>
- Plante, T., Stanley, L., Nejadpak, A., & Yang, C. X. (2016). Rotating machine fault detection using principal component analysis of vibration signal.
- Pourebahimi, A., & Mahmoodi, M. (2010). Status Detection and Fault Diagnosing of Rotatory Machinery by Vibration Analysis Using Data Mining, (Icctd), 131–135.
- Qiu, H., Lee, J., Lin, J., & Yu, G. (2006). Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics. *Journal of Sound and Vibration*, *289*(4–5), 1066–1090. <https://doi.org/10.1016/j.jsv.2005.03.007>
- Rafiee, J., Rafiee, M. A., & Tse, P. W. (2010). Application of mother wavelet functions for automatic gear and bearing fault diagnosis. *Expert Systems with Applications*, *37*(6), 4568–4579. <https://doi.org/10.1016/j.eswa.2009.12.051>
- Rajeswari, C., Sathiyabhama, B., Devendiran, S., & Manivannan, K. (2014). Bearing fault diagnosis using wavelet packet transform, hybrid PSO and support vector machine. *Procedia Engineering*, *97*, 1772–1783. <https://doi.org/10.1016/j.proeng.2014.12.329>
- Rashid, M. M., Amar, M., Gondal, I., & Kamruzzaman, J. (2016). A data mining approach for machine fault diagnosis based on associated frequency patterns. *Applied Intelligence*, *45*(3), 638–651. <https://doi.org/10.1007/s10489-016-0781-3>
- Roemer Vlasveld. (2013). Roemer's blog. Retrieved January 1, 2017, from <http://rvlasveld.github.io/blog/2013/07/12/introduction-to-one-class-support-vector-machines/>
- Safizadeh, M. S., Lakis, A. A., & Thomas, M. (2000). Using Short-Time Fourier Transform in Machinery Diagnosis. *International Journal of COMADE*, *3*(1), 1–14.
- Santos, M. F., & Azevedo, C. (2005). *Data Mining-Descoberta de conhecimento em bases de dados*.
- Sawalhi, N., & Randall, R. B. (2005). Spectral kurtosis optimization for rolling element bearings. *Proceedings - 8th International Symposium on Signal Processing and Its Applications, ISSPA 2005*, *2*, 839–842. <https://doi.org/10.1109/ISSPA.2005.1581069>
- Selvakumar, J., & Edwin Prakash, A. (2016). Data Acquisition System and Signal Processing Technique for Bearing Fault Analysis. *Indian Journal of Science and Technology*, *9*(31). <https://doi.org/10.17485/ijst/2016/v9i31/92769>
- Shen, Z., He, Z., Chen, X., Sun, C., & Liu, Z. (2012). A monotonic degradation assessment index of rolling bearings using fuzzy support vector data description and running time. *Sensors (Basel, Switzerland)*, *12*(8), 10109–35. <https://doi.org/10.3390/s120810109>
- Soualhi, A., Medjaher, K., & Zerhouni, N. (2015). Bearing Health Monitoring Based on Hilbert – Huang Transform , Support Vector Machine , and Regression, *64*(1), 52–62. <https://doi.org/10.1109/TIM.2014.2330494>
- Suganya, R., & Shanthi, R. (2012). Fuzzy C- Means Algorithm- A Review. *International Journal*, *2*(11), 1–3.
- Tian, J., Morillo, C., Azarian, M. H., & Pecht, M. (2016). Kurtosis-Based Feature Extraction Coupled With K -Nearest Neighbor Distance Analysis. *IEEE Trans. Ind. Electron.*, *63*(3), 1793–1803. <https://doi.org/10.1109/TIE.2015.2509913>
- Units, M., & Influences, O. E. (1982). Measuring Vibration.

- University, C. W. R. (2017). Bearing Data Center Website. Retrieved January 1, 2017, from <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>
- Valens, C. (1999). A Really Friendly Guide to Wavelets.
- Vujovi, I. (2015). Analysis of the vibration signal using time-frequency methods, 3, 23–35.
- Wang, Y., & Markert, R. (2015). Spectral kurtosis for fault detection , diagnosis and prognostics of rotating machines : A review with applications. *Mechanical Systems and Signal Processing*, 66–67(June), 679–698. <https://doi.org/10.1016/j.ymssp.2015.04.039>
- Xia, Z., Xia, S., Wan, L., & Cai, S. (2012). Spectral regression based fault feature extraction for bearing accelerometer sensor signals. *Sensors (Switzerland)*, 12(10), 13694–13719. <https://doi.org/10.3390/s121013694>
- Yan, R., Member, S., Gao, R. X., & Member, S. (2006). Hilbert – Huang Transform-Based Vibration Signal Analysis for Machine Health Monitoring, 55(6), 2320–2329.
- Yang, Q., Guo, J., Zhang, D., & Liu, C. (2011). Fault diagnosis based on fuzzy C-means algorithm of the optimal number of clusters and probabilistic neural netWork. *International Journal of Intelligent Engineering & Systems*, 4(2), 51–59.
- Yiakopoulos, C. T., Gryllias, K. C., & Antoniadis, I. A. (2011). Rolling element bearing fault detection in industrial environments based on a K-means clustering approach. *Expert Systems with Applications*, 38(3), 2888–2911. <https://doi.org/10.1016/j.eswa.2010.08.083>
- Zhang, Z., Wang, Y., & Wang, K. (2013). Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *Journal of Intelligent Manufacturing*, 24(6), 1213–1227. <https://doi.org/10.1007/s10845-012-0657-2>
- Zhao, M., Jin, X., Zhang, Z., & Li, B. (2014). Fault diagnosis of rolling element bearings via discriminative subspace learning: Visualization and classification. *Expert Systems with Applications*, 41(7), 3391–3401. <https://doi.org/10.1016/j.eswa.2013.11.026>
- Zhigang, Z., Hongmao, Q., & Ning, S. (2016). Bearing Fault Diagnosis Based on Wavelet Packet Decomposition Energy Distribution and Gray Similar Correlation Degree, 9(2), 109–120.
- Zhou, H., Chen, J., Dong, G., & Wang, R. (2016). Detection and diagnosis of bearing faults using shift-invariant dictionary learning and hidden Markov model. *Mechanical Systems and Signal Processing*, 72–73, 65–79. <https://doi.org/10.1016/j.ymssp.2015.11.022>
- Zhou, J., Guo, H., Zhang, L., Xu, Q., & Li, H. (2016). Lifting Wavelet Packet Symbolic Entropy and SVDD, 2016.
- Zhu, X., Zhang, Y., & Zhu, Y. (2013). Bearing performance degradation assessment based on the rough support vector data description. *Mechanical Systems and Signal Processing*, 34(1–2), 203–217. <https://doi.org/10.1016/j.ymssp.2012.08.008>