

Aplicação de Machine Learning e Descritores Estatísticos para Diagnóstico de Falhas em Rolamentos de Motores Elétricos Industriais

Lucas de Oliveira Soares* Luiz Alberto Pinto*
Diego Assereuy Lobão*

* Programa de Pós-Graduação em Engenharia de Controle e Automação – Propecaut, Instituto Federal do Espírito Santo, Serra, Brasil (e-mails: lucasosoares@hotmail.com, pintoluizalberto@gmail.com, diego.lobao@gmail.com)

Abstract: This paper presents a performance analyzes of the classifiers k-Nearest Neighbor, Support Vector Machine, Decision Tree and Random Forest, in classifying vibration signals acquired from normal bearings and bearings in which artificial failures were introduced. Two approaches were considered for descriptors extraction. The first one uses Principal Component Analysis (PCA) to extract the main components from the original signals, which are used as descriptors for modeling. The second one obtains the models from the signals statistical descriptors in the time domain. The data set used in the tests is the CWRU Bearing Dataset, and the results show the statistical descriptors present good class separability favoring the performance of the classification models, which obtain rates greater than 99% of accuracy in the Decision Tree and Random Forest classifiers.

Resumo: Esse artigo faz uma análise do desempenho dos classificadores *k-Nearest Neighbour*, *Support Vector Machine*, *Decision Tree* e *Random Forest*, para a classificação de sinais de vibração adquiridos a partir de rolamentos normais e de rolamentos nos quais foram introduzidas falhas artificiais. Para a extração de descritores foram consideradas duas abordagens. A primeira utiliza Análise de Componente Principal (*PCA*) para extração das componentes principais dos sinais originais, que são utilizadas como descritores para a modelagem. A segunda obtém os modelos a partir de descritores estatísticos dos sinais no domínio do tempo. O conjunto de dados utilizados nos testes é o *CWRU Bearing Dataset*, e os resultados mostram que os descritores estatísticos apresentam boa separabilidade das classes favorecendo o desempenho dos modelos de classificação, que obtém taxas superiores a 99% de acurácia nos classificadores Árvore de Decisão e Floresta Aleatória.

Keywords: bearing fault diagnosis; pattern recognition; statistical descriptors; feature extraction; principal component analysis.

Palavras-chaves: diagnóstico de falhas em rolamentos; reconhecimento de padrões; descritores estatísticos; extração de características; análise de componentes principais.

1. INTRODUÇÃO

Ambientes industriais são constituídos por diferentes tipos de equipamentos que fazem parte de sistemas de produção complexos. Entre os diversos tipos de equipamentos que operam nesses ambientes, um dos mais importantes e mais sujeitos a falhas são as máquinas rotativas. Dessa forma, a utilização de técnicas de manutenção eficientes são importantes para que esses equipamentos estejam sempre em condições de produção. As estratégias mais eficientes de manutenção, consistem no monitoramento automático contínuo das condições de funcionamento das máquinas rotativas, através da coleta e análise de sinais por eles produzidos quando em funcionamento. A implantação de tal estratégia em um ambiente industrial pode reduzir os custos das intervenções e os tempos de parada da produção.

A utilização de técnicas de monitoramento automático das condições de funcionamento de equipamentos através da análise dos sinais do processo, podem auxiliar na identificação, no diagnóstico e na eliminação das falhas em máquinas rotativas. Zhang et al. (2017), afirmam que modelos de diagnóstico de falhas baseados em métodos orientados a dados são vantajosos, pois podem fornecer diagnósticos precisos, mesmo sem o conhecimento do modelo do processo. Nesses métodos, os diagnósticos são elaborados a partir de dados que, de forma geral, são facilmente obtidos por sensores.

Devido a sua natureza dinâmica, quando em funcionamento as máquinas rotativas são submetidas a ação de forças centrífugas, forças de atrito e forças resultantes da vibração das partes móveis ou de folgas no conjunto. Devido a esse sistema de forças atuando sobre o equipamento, as partes móveis sofrem desgastes operacionais, que com

o tempo podem resultar no aparecimento de falhas e em quebras intempestivas.

Nesse contexto, os mancais de rolamento se tornam as partes mais importantes das máquinas rotativas. Sob condições adequadas de operação e conservação os mancais de rolamento podem reduzir os efeitos das ações das referidas forças sobre o conjunto, prevenindo o aparecimento de alguns tipos de falhas. De acordo com Zhang et al. (2019), que investigou a probabilidade de ocorrência de falhas em máquinas rotativas, problemas em rolamentos são responsáveis por cerca de 30% a 40% de todas as falhas nessas máquinas. Como, geralmente, condições anormais de funcionamento de rolamentos se manifestam na forma de vibração, o monitoramento dessa variável é uma das principais ferramentas para acompanhamento das condições de funcionamento das máquinas.

Em função da relevância deste tema, constatamos que rolamentos são componentes básicos essenciais para uma máquina rotativa e isto motiva o estudo da detecção antecipada de problemas para que não haja impactos nos processos produtivos onde eles estão presentes. Com a crescente demanda de monitoramento e a grande quantidade de informações de dados de processos industriais, a utilização de aprendizagem de máquina está ganhando cada vez mais atenção pois, por meio destas técnicas, somos capazes de retirar informações úteis de dados históricos (Chen et al., 2019).

Este artigo aborda a utilização de técnicas de *Machine Learning*, *K-Nearest Neighbour*, *Support Vector Machine*, *Decision Tree* e *Random Forest*, para classificação de falhas extraídas de sinais de vibração produzidos por rolamentos de motores. O conjunto de dados utilizado é o da *Case Western Reserve University - CWRU*, disponível abertamente para estudos na área acadêmica. O objetivo é apresentar o desempenho destas ferramentas de classificação em função dos descritores extraídos do sinal vibracional e dos modos de operação do equipamento em função das cargas de trabalho.

Na sequência este trabalho está constituído pelas seguintes seções: a Seção 2 contém a revisão do estado da arte sobre diagnóstico de falhas; a Seção 3 apresenta uma descrição sucinta dos algoritmos de classificação empregados; a Seção 4 traz uma descrição das etapas executadas na realização da pesquisa; a Seção 5 reúne os resultados obtidos e a Seção 6 apresenta as conclusões do trabalho.

2. REVISÃO BIBLIOGRÁFICA

Sistemas de diagnóstico têm como base a detecção e a classificação das condições anormais de funcionamento. Devido a sua importância no setor industrial, diversos pesquisadores têm se dedicado a estudar a aplicação de técnicas de *Machine Learning* a elaboração de diagnóstico de falhas em sistemas industriais.

Rauber et al. (2015) estudaram a classificação de sinais de vibração utilizando os algoritmos *k-Nearest Neighbors (k-NN)*, *Support Vector Machine (SVM)* e Redes Neurais Artificiais *Multilayer Perceptron (MLP)* para o diagnóstico de falhas em rolamentos. Nesse trabalho, são comparados os desempenhos dos classificadores quando implementados com a utilização de descritores obtidos com base em

modelos estatísticos, *Wavelet Package Analysis* e análise complexa de envelope. Para a redução da dimensionalidade os autores utilizaram Análise de Componentes Principais - PCA e seleção de características, sendo que os melhores resultados foram obtidos por métodos de seleção de características. O modelo construído pela combinação de *Wavelet Package* e o *k-NN* apresentou 99,82% de acurácia, sendo este o melhor valor encontrado.

Nayana and Geethanjali (2017), estudaram a modelagem de sinais de vibração para a detecção de falhas em rolamentos, através de parâmetros estatísticos no domínio do tempo. Os descritores estatísticos utilizados para a construção dos modelos foram o comprimento de forma de onda (*WL*), alterações no sinal de inclinação (*SSC*), integral de sinal simples (*SSI*), amplitude de Wilson (*WAMP*), valor absoluto médio (*MAV*) e cruzamento zero (*ZC*) sendo estes comparados com os descritores convencionais tais como desvio padrão (*STD*), variância (*VAR*), assimetria (*SKW*), valor médio (*Mean*), raiz média quadrática (*RMS*) e curtose (*KURT*). O estudo realizou a análise dos descritores citados, com uso dos classificadores implementados pelos algoritmos *Naive Bayes (NB)*, *SVM* e *Linear Discriminant Analysis (LDA)*. A validação dos modelos foi realizada pela técnica de *k-Fold Cross Validation* com 10 folds. Na comparação final do método proposto, com uso dos descritores *WL*, *SSC*, *SSI*, *WAMP*, *MAV* e *ZC* contra os descritores convencionais no domínio do tempo, todos os classificadores atingiram 100% de acurácia, independente da carga do motor do experimento. Trabalhos realizados por outros autores utilizando o mesmo conjunto de dados, porém com descritores diferentes (*STD*, *VAR*, *SKW*, *Mean*, *RMS* e *KURT*), obtiveram resultados entre 73% e 99%.

Em Tahir et al. (2017) os autores propuseram uma abordagem para melhoria da acurácia na classificação de falhas em rolamentos de esferas utilizando preditores extraídos dos sinais de vibração no domínio do tempo. O conjunto de dados utilizado utilizados foi obtido em um simulador de falhas na universidade de Curtin-Austrália. Além da classe relacionada ao funcionamento normal, foram registradas medições associadas à problemas nas pistas interna e externa do rolamento. Para a classificação, dez descritores estatísticos no domínio do tempo foram extraídos dos sinais de vibração, que constituíram os dados de entrada para os classificadores *SVM*, *BayesNet*, *Decision Table* e *Decision Tree (DT)*. Segundo os autores, a abordagem proposta apresentou vantagens substanciais em relação ao custo computacional, pois a utilização dos referidos descritores resultou em modelos eficientes para os sinais de vibração, permitindo a representação dos mesmos por uma pequena quantidade de descritores estatísticos, sem perda substancial de informação. Os modelos com os descritores *RMS* e *variância*, em todos os casos atingiram o valor de 100%. Os classificadores com pior desempenho foram aqueles obtidos com descritores extraídos por valores médios e mediana no algoritmo *DT*, com valores de 73% e 84% respectivamente.

No trabalho de Dey et al. (2019) é proposta a utilização da técnica de autocorrelação com os sinais no domínio do tempo para a detecção de falhas em rolamentos de esferas. Para a obtenção dos modelos os autores extraíram os descritores estatísticos, valor máximo, desvio padrão, variância e valores de pico a pico dos correlogramas construídos.

Na etapa de classificação, foram utilizados os algoritmos *Random Forest (RF)* e *k-NN*, e a metodologia consistiu na divisão dos processos em três problemas binários e um problema multi-classes. Ambos os classificadores apresentaram acurácia de 100%, o que segundo os autores superou os resultados até então relatados.

Trabalhos recentes mostram uma tendência no uso de técnicas de redes neurais e algoritmos genéticos para diagnóstico de falhas em rolamentos. Toma et al. (2020) apresenta uma combinação de técnicas que envolve a utilização de métodos estatísticos, algoritmo genético (AG) e técnicas de aprendizado de máquina para o diagnóstico de falhas em motores. Os descritores são extraídos dos sinais de corrente elétrica com a utilização de AG. Modelos de classificação baseados em *k-NN*, *DT* e *RF* são treinados e aplicados a identificação de falhas nos rolamentos do equipamento. Os autores relatam taxas de acerto superiores a 97%. Agrawal and Jayaswal (2019) fizeram um estudo comparativo entre uma rede neural artificial (RNA) e *SVM* para o diagnóstico e classificação de falhas em rolamento utilizando wavelets contínuas e entropia. Os resultados indicaram o melhor desempenho do *SVM*, que obteve acurácia de 100%, contra 96% das redes neurais.

Em Zhou et al. (2019) é utilizada uma Rede Neural Convocional (RNC) para classificação do estado de degradação dos rolamentos. O procedimento consistiu de duas etapas. Na etapa não supervisionada, descritores dos sinais no domínio do tempo são extraídos, e utilizados como entrada de um algoritmo de clusterização *k-Means*, onde classes são constituídas. Na etapa supervisionada os dados categorizados são fornecidos como entrada para a RNC. A acurácia obtida no treino e teste foram, respectivamente, 98.89% e 98.58%.

3. REFERENCIAL TEÓRICO

Embora a maioria dos trabalhos recentes abordando o tema do diagnóstico de falhas em sistemas industriais utilizem métodos baseados em *Deep Learning*, os algoritmos de *Machine Learning* convencionais ainda são utilizados para essa finalidade. A seguir é apresentada uma breve descrição dos algoritmos *Support Vector Machine*, *K-Nearest Neighbour*, *Decision Tree* e *Random Forest*, aplicados para o treinamento dos modelos.

3.1 Support Vector Machine - SVM

Support Vector Machine são algoritmos de classificação supervisionada, lineares e binários, que procuram o hiperplano de separação entre as classes (Figura 1), de maneira a maximizar a distância entre as amostras mais próximas (Cortes and Vapnik, 1995). A aplicação do *SVM* à problemas não lineares pode ser possível com a utilização de funções *kernel*, que projetam o espaço amostral sobre espaços de maior dimensão onde os dados são linearmente separáveis. Exemplos de funções *kernel* que podem ser utilizadas são, quadrática, gaussiana, polinomial. Embora seja, em sua forma convencional, um algoritmo binário, que suporta apenas problemas com duas classes, a utilização de métodos multiclasse torna o *SVM* aplicável a problemas com mais de duas classes.

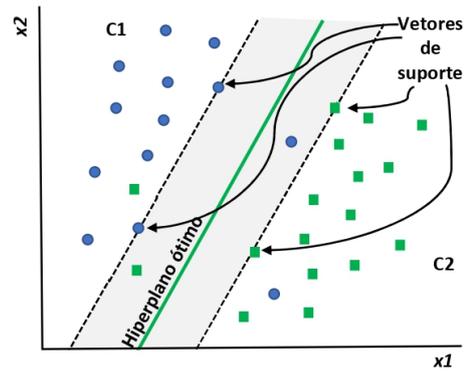


Figura 1. Exemplo ilustrativo do algoritmo SVM.

3.2 k-Nearest Neighbour - k-NN

Classificadores implementados com base no algoritmo *k-NN* classificam amostras desconhecidas com base na distância entre essas e as amostras atribuídas às classes envolvidas no problema (Duda et al., 2001). Por ser de fácil implementação, esses classificadores são amplamente utilizados em problemas de classificação. Para sua implementação devem ser definidos, o número de vizinhos (*k*), da métrica para o cálculo da distância entre as amostras e a regra de decisão para atribuição de uma classe a uma amostra desconhecida. Uma regra de decisão comumente utilizada é atribuir à amostra desconhecida a classe da maioria das *k* amostras mais próximas. A Figura 2 apresenta uma ilustração de classificação utilizando o algoritmo 4-*NN* em um problema de três classes. Um conjunto de dados cujos objetos pertencem à *C1*, *C2* ou *C3* é representado pelos descritores x_1 e x_2 . O objeto *AD*, cuja classe é desconhecida, será atribuído a uma das classes do problema. Com $k = 4$, *AD* será atribuído a *C3*, pois a maioria dos $k = 4$ objetos mais próximos de *AD* pertencem a *C3*.

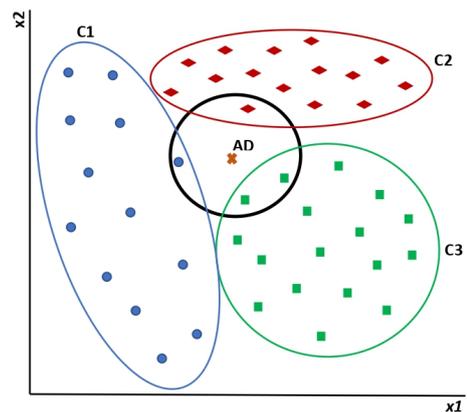


Figura 2. Exemplo ilustrativo do algoritmo *k-NN* com $k=4$.

3.3 Decision Tree - DT

A *DT* é um método de classificação que utiliza a estratégia dividir-para-conquistar, cujo processo de decisão para atribuição de uma classe é constituído por múltiplos estágios (Webb, 2003). Sua estrutura hierárquica em forma de árvore invertida o torna um classificador de simples

implementação. As *DT* são constituídas por nós e ramos. As regras de decisão estão associadas aos nós, onde testes são realizados para a definição do percurso ao longo da árvore. Os ramos representam os possíveis caminhos a serem percorridos de acordo com a decisão tomada no nó precedente. Decisões sucessivas são tomadas até que um nó terminal seja alcançado, onde uma classe é atribuída à amostra desconhecida (Dougherty, 2013). A implementação da *DT* requer a definição prévia da profundidade da árvore e do critério de decisão.

3.4 Random Forest - RF

A *RF* é um algoritmo de aprendizado de máquina supervisionado, que pode ser aplicado à problemas de classificação e regressão. Uma *RF* é constituída por um conjunto de Árvores de Decisão, cujo algoritmo de treinamento mais comumente utilizado é o bagging. Como utiliza combinação das árvores de decisão, a Floresta Aleatória, de forma geral, é mais estável e apresenta melhor desempenho quando comparada a outros algoritmos de *Machine Learning*. A premissa sobre a qual se baseia a *RF* é que, um vetor aleatório v_k para a k -ésima árvore, independente dos vetores aleatórios passados $v_1, v_2, v_3, \dots, v_{k-1}$, mas com a mesma distribuição será criado. Com estes vetores aleatórios gerados, em um processo de votação envolvendo as k árvores, uma classe será atribuída à amostra desconhecida (Breiman, 2001).

4. METODOLOGIA

Nessa seção é feita uma descrição detalhada da metodologia que orientou a realização do trabalho. O conjunto de dados utilizado para a realização dos testes é apresentado, bem como os descritores extraídos a partir dos sinais. São descritas as etapas de pré-processamento e os ajustes dos parâmetros utilizados para a configuração dos classificadores.

4.1 Descrição do conjunto de dados

O conjunto de dados de vibração utilizado é um dos mais populares no meio acadêmico para testes com modelos de *Machine Learning* e foi desenvolvido pela *Case Western Reserve University - CWRU*. A Figura 3 ilustra a bancada experimental utilizada para a aquisição dos sinais. Um motor com potência nominal de 2hp foi posto em funcionamento com cargas de 0hp, 1hp, 2hp e 3hp. Os rolamentos do lado acionado (DE – Drive End) e do lado não acionado (NDE - Non Drive End) foram danificados artificialmente utilizando eletro-descargas criando erosões com diâmetros de 0,1778 mm a 1,016 mm. Os sinais de vibração foram coletados por acelerômetros a frequências de amostragem de 12kHz e 48kHz. O conjunto de dados original é constituído por seis classes colhidas do rolamento DE. Além da condição de funcionamento normal (HDE), foram simuladas falha na pista interna (IR), falhas nas esferas (BA) e na pista externa em três posições diferentes, centralizada (ORCENT), ortogonal (ORORT) e oposta (OROPP).

Os arquivos no formato Matlab (*.mat) com os sinais coletados nos testes podem ser encontrados no link <https://>

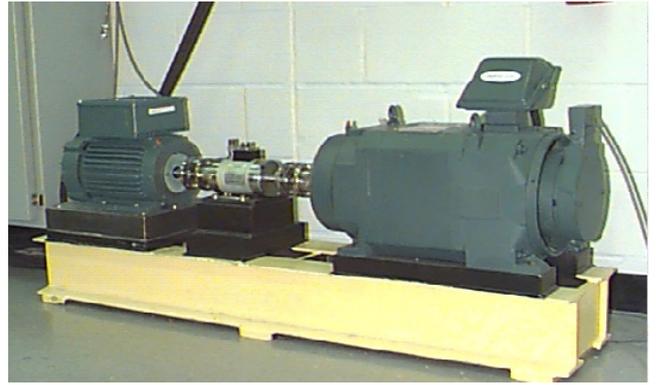


Figura 3. Conjunto experimental para a coleta do sinal de vibração dos mancais do conjunto de dados CWRU.

csegroups.case.edu/bearingdatacenter/home, e estão organizados de acordo com a Tabela 1. Cada arquivo se refere a uma classe e corresponde a uma potência de carga utilizada para o teste. Por exemplo, o arquivo 119 se refere a falha BA com o motor com carga de 1hp. Para cada classe estão disponíveis quatro arquivos referentes aos testes nas potências de 0hp, 1hp, 2hp e 3hp.

Tabela 1. Organização dos arquivos com os sinais originais no site da CWRU.

Condição rolamento	Nome do arquivo (*.mat)			
	0hp	1hp	2hp	3hp
HDE	97	98	99	100
IR	105	106	107	108
BA	118	119	120	121
ORCENT	130	131	132	133
ORORT	144	145	146	147
OROPP	156	158	159	160

Para os testes realizados nesse trabalho foram utilizadas somente as medições obtidas no rolamento do lado acionado (DE), com diâmetro de falha de 0,1778 mm em todas as potências as quais o motor foi submetido. A Tabela 2 mostra a quantidade de pontos amostrais dos sinais coletados para a classe normal e para as falhas simuladas em cada potência.

Tabela 2. Quantidade de amostras por falha e potência dos sinais colhidos no rolamento DE.

Condição rolamento	Diâmetro Falha (mm)	Número de pontos amostrais			
		0hp	1hp	2hp	3hp
HDE	-	243938	483903	485063	485643
IR	0,1778	121265	121991	122136	122917
BA	0,1778	122571	121410	121556	121556
ORCENT	0,1778	121991	122426	121410	122571
ORORT	0,1778	122281	121846	121556	122281
OROPP	0,1778	122281	121991	122281	122136

4.2 Segmentação dos Sinais

Os sinais originais de vibração colhidos no experimento (Tabela 2), para cada condição de funcionamento a que o rolamento DE foi submetido nos testes, foram segmentados em amostras de 2000 pontos que constituíram as classes. A coluna 3 da Tabela 3 mostra as quantidades de amostras e atributos para cada classe para cada potência do motor.

4.3 Extração de descritores

Os modelos de classificação foram treinados com descritores extraídos por duas abordagens diferentes.

Análise de componentes principais: inicialmente, as componentes principais da matriz de dados originais, extraídas com a aplicação da PCA foram utilizadas como descritores para a obtenção dos modelos. Nesse caso, deseja-se avaliar o desempenho dos modelos a partir de um formato reduzido dos dados diretamente obtidos do processo. Para garantir que a quase totalidade das informações relevantes se mantenham na forma reduzida dos dados originais, foi assegurado que, no mínimo 95% da variabilidade contida na base de dados original seja preservada na forma reduzida. A coluna 4 da Tabela 3, mostra as dimensões das matrizes de descritores por classes, obtidos com a aplicação da PCA sobre o conjunto original. Importante destacar que a extração das componentes principais utilizadas para a modelagem foi realizada posteriormente ao particionamento dos conjuntos de treino e teste, não havendo portanto influência do conjunto de treino sobre o de teste.

Descritores estatísticos: vários trabalhos apresentados na Seção 2, mostraram o bom desempenho de modelos treinados através de descritores estatísticos no domínio do tempo. As Equações de 1 a 5 mostram as definições matemáticas dos descritores utilizados nessa pesquisa.

$$Média(\mu) = \frac{1}{N} \sum_{i=1}^N s(i) \quad (1)$$

$$Variância(\sigma^2) = \frac{1}{N} \sum_{i=1}^N (s(i) - \mu)^2 \quad (2)$$

$$Assimetria = \frac{1}{N} \sum_{i=1}^N \left(\frac{s(i) - \mu}{\sigma} \right)^3 \quad (3)$$

$$Curtose = \frac{1}{N} \sum_{i=1}^N \left(\frac{s(i) - \mu}{\sigma} \right)^4 \quad (4)$$

$$Fator\ de\ Crista = \frac{máx(|s|)}{RMS} \quad (5)$$

Em cada equação, o “s” representa as amostras individuais do conjunto de dados em cada teste. A coluna 5 da Tabela 3 apresenta as dimensões das matrizes de descritores estatísticos por classe, após a aplicação das Equações 1 a 5. Após a extração dos descritores estatísticos, as matrizes das classes foram concatenadas formando matrizes de 420x5 para a potência de 0hp, e 540x5 para as demais potências.

4.4 Análise do espaço amostral dos conjuntos de descritores derivados da PCA e pelo método estatístico

A obtenção de modelos com boa capacidade de separação das classes, além do bom desempenho dos algoritmos, depende da extração de descritores que, de fato, forneçam uma boa representação das amostras no espaço amostral.

Tabela 3. Dimensões das matrizes por classe e por potência com os sinais originais segmentados (coluna 3), após aplicação da PCA (coluna 4) e após extração de descritores estatísticos (coluna 5).

Potência	Classes	Matrizes de atributos		
		Original	PCA	Estatístico
0hp	HDE	120x2000	120x104	120x5
	IR	60x2000	60x104	60x5
	BA	60x2000	60x104	60x5
	ORCENT	60x2000	60x104	60x5
	ORORT	60x2000	60x104	60x5
1hp	OROPP	60x2000	60x104	60x5
	HDE	240x2000	240x115	240x5
	IR	60x2000	60x115	60x5
	BA	60x2000	60x115	60x5
	ORCENT	60x2000	60x115	60x5
2hp	ORORT	60x2000	60x115	60x5
	OROPP	60x2000	60x115	60x5
	HDE	240x2000	240x114	240x5
	IR	60x2000	60x114	60x5
	BA	60x2000	60x114	60x5
3hp	ORCENT	60x2000	60x114	60x5
	ORORT	60x2000	60x114	60x5
	OROPP	60x2000	60x114	60x5
	HDE	240x2000	240x121	240x5
	IR	60x2000	60x121	60x5
	BA	60x2000	60x121	60x5
	ORCENT	60x2000	60x121	60x5
	ORORT	60x2000	60x121	60x5
	OROPP	60x2000	60x121	60x5

As Figuras 4 e 5 ilustram os espaços amostrais dos conjuntos de descritores obtidos pela PCA e pelos descritores estatísticos, respectivamente, para as quatro potências de carga.

A Figura 4 ilustra o espaço amostral representado pelas duas primeiras componentes principais após aplicação da PCA. Pode ser notado que, para as quatro potências de carga, há uma superposição acentuada das amostras das classes. Esse arranjo das classes pode representar uma maior dificuldade para a obtenção de classificadores com boa capacidade de separabilidade e de predição a partir de descritores derivados pela aplicação da PCA no conjunto de dados originais.

A Figura 5 ilustra o espaço amostral representado pelos descritores estatísticos média x variância. Como pode ser observado, a utilização dos descritores estatísticos para representação dos sinais de vibração, comparativamente aos descritores oriundos da PCA, melhoram significativamente a separabilidade das classes, o que favorece a obtenção de classificadores com boa capacidade de predição.

4.5 Treinamento dos modelos

Os modelos de classificação foram treinados com os algoritmos *k-NN*, *SVM*, *DT* e *RF*. Os critérios de divisão utilizados para implementação dos classificadores *DT* foram o *gini index* e o *cross entropy*, sendo em todos os casos utilizadas árvores com profundidade 10. No algoritmo *k-NN* foi utilizada a métrica de distância euclidiana e variados os números de vizinhos em 1, 3 e 5. No classificador *SVM* foram testados os *kernels linear*, *gaussian* e *polynomial*. Com a *RF* foram testados os modelos variando o número

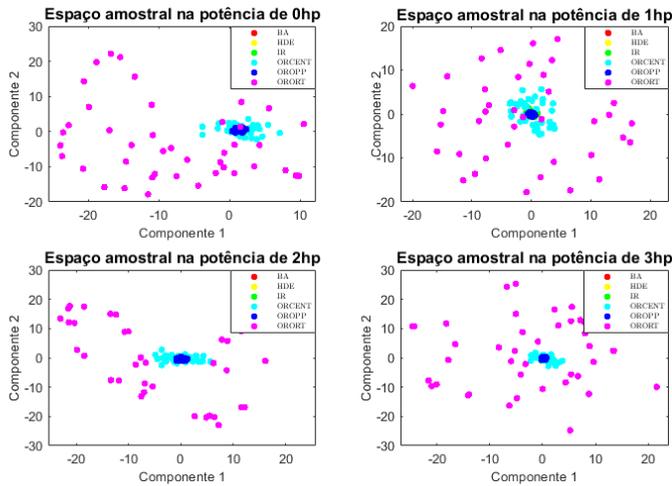


Figura 4. Espaços amostrais com duas componentes principais para potências de 0hp, 1hp, 2hp e 3hp.

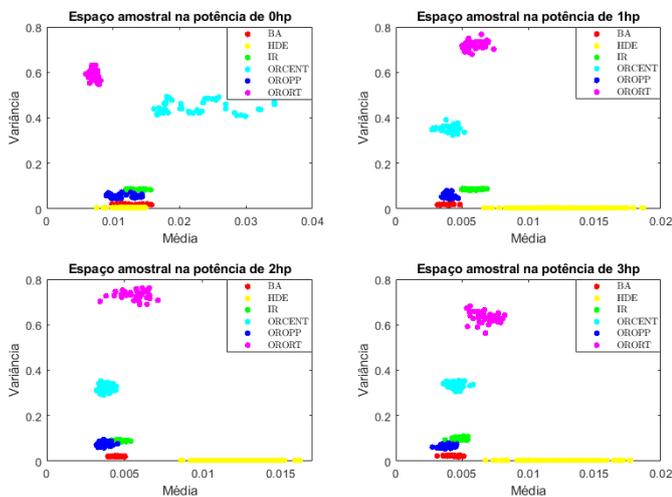


Figura 5. Espaços amostrais com atributos estatísticos para potências de 0hp, 1hp, 2hp e 3hp.

árvores, sendo 5,10, 15 e 20 para os descritores estatísticos e 10, 20, 50, 100 e 200 para PCA. O conjunto de dados foi dividido em dados de treinamento, com 70% das amostras, e dados de testes com os 30% restantes. O desempenho dos modelos foi avaliado pela técnica k-Fold Cross Validation com $k=10$, e foi utilizada a métrica de acurácia. Todos os testes foram realizados utilizando o Matlab®, versão R2018b.

5. RESULTADOS

Esta seção apresenta os resultados de classificação obtidos durante as etapas de treinamento e testes conforme a metodologia apresentada anteriormente. Na sequência, os resultados serão apresentados (i) utilizando os descritores extraídos através da PCA e (ii) através dos dados estatísticos.

5.1 Resultados utilizando descritores através da PCA

As Tabelas 4, 5, 6 e 7 apresentam, respectivamente, os resultados obtidos pelos classificadores *DT*, *k-NN*, *SVM*, e *RF*. A métrica utilizada para medir o desempenho dos

classificadores durante as etapas de treinamento e testes foi a acurácia. Conforme pode ser observado nas tabelas, o desempenho dos classificadores durante as etapas de treinamento e testes apresentou baixa performance. O melhor resultado é obtido utilizando o classificador *SVM* (Tabela 6), com acurácia de 85,70% utilizando o *kernel gaussian* sem carga no motor (0 hp) e acurácia de 88,90% com cargas no motor variando de 1 a 3 (hp). Contudo, ao se comparar os resultados dos treinos e dos testes para esse classificador, nota-se que na maioria dos testes ocorreu um problema de *underfitting*, isto é, os modelos apresentaram baixo desempenho no treino e desempenho significativamente melhor nos testes. O oposto disso ocorreu com a maioria dos modelos obtidos por *DT*. Nesse caso, exceto para a potência de 3 hp, todos os modelos sofreram *overfitting*, isto é, o desempenho no treino foi significativamente melhor que no teste, embora os resultados no treino não tenham sido bons. Conforme visto anteriormente na Figura 4, o espaço amostral representado desse conjunto de dados apresenta sobreposições interclasses e grande dispersão na classe *ORORT*, isso explica o baixo desempenho dos modelos devido a pouca separabilidade das classes nos descritores utilizados.

Tabela 4. Resultados com classificador *DT*

Potência[hp]	Parâmetro	Acurácia [%]	
		Treino	Teste
0	Entropy	61,22	14,30
	Gini	61,22	23,00
1	Entropy	57,67	17,90
	Gini	61,38	14,20
2	Entropy	74,60	24,10
	Gini	71,96	23,50
3	Entropy	67,72	23,50
	Gini	65,34	19,10

Tabela 5. Resultados com classificador *k-NN*

Potência[hp]	Parâmetro	Acurácia [%]	
		Treino	Teste
0	k=1	75,51	71,40
	k=3	63,95	73,00
	k=5	58,84	76,20
1	k=1	66,93	69,70
	k=3	65,34	63,60
	k=5	62,17	59,30
2	k=1	76,46	67,90
	k=3	71,16	61,70
	k=5	68,78	60,10
3	k=1	74,87	67,30
	k=3	68,25	62,30
	k=5	64,29	59,90

5.2 Resultados utilizando descritores estatísticos

As Tabelas 8, 9, 10 e 11 apresentam, respectivamente, os resultados obtidos pelos classificadores *DT*, *k-NN*, *SVM*, e *RF*. A mesma metodologia utilizada anteriormente, com os descritores baseados na PCA, foi utilizada nas etapas de treinamento e testes com os dados estatísticos. De forma geral, todos os modelos obtiveram bom desempenho nas etapas de treinamento e testes, atingindo valores de acurácia em 100%.

Nos resultados utilizando *DT* (Tabela 8), é possível verificar que a regra de divisão utilizando *entropy* foi ligeiramente superior ao método utilizando índice *gini*. Para o

Tabela 6. Resultados com classificador SVM

Potência[hp]	Parâmetro	Acurácia [%]	
		Treino	Teste
0	Gaussian	42,18	85,70
	linear	38,44	82,50
	Polynomial	30,95	69,00
1	Gaussian	55,29	88,90
	linear	48,94	75,30
2	Polynomial	46,83	55,60
	Gaussian	55,03	88,90
3	linear	59,26	65,40
	Polynomial	52,91	55,6
3	Gaussian	55,56	88,9
	linear	56,88	77,20
	Polynomial	46,83	55,6

Tabela 7. Resultados com classificador RF

Potência[hp]	Parâmetro	Acurácia [%]	
		Treino	Teste
0	10 trees	63,61	77,00
	20 trees	70,75	74,60
	50 trees	80,95	69,00
	100 trees	85,71	68,30
	200 trees	87,76	69,80
1	10 trees	67,99	78,40
	20 trees	73,28	79,00
	50 trees	78,04	79,60
	100 trees	80,95	82,70
	200 trees	85,98	79,60
2	10 trees	68,78	77,20
	20 trees	78,31	77,20
	50 trees	83,86	73,50
	100 trees	88,62	74,10
	200 trees	91,80	69,80
3	10 trees	70,63	74,10
	20 trees	77,51	72,80
	50 trees	84,66	74,70
	100 trees	86,24	71,00
	200 trees	87,30	74,10

classificador k -NN (Tabela 9), em todas as configurações utilizadas para as diferentes cargas de potência do motor, foi possível obter acurácia acima de 91,30% sendo que o melhor resultado foi obtido utilizando o valor de $k = 3$, o qual obteve uma acurácia de 100% nas etapas de treinamento e testes. O classificador SVM (Tabela 10), de forma geral, obteve os melhores resultados utilizando o *kernel* do tipo *polynomial*, no qual em todas as configurações obteve acurácia em 100% nas etapas de treinamento e testes. Os resultados de classificação utilizando o modelo RF (Tabela 11) obtiveram desempenhos acima de 98,40% de acurácia, sendo que, a partir de uma coleção igual ou superior a dez árvores, em todas as configurações foi possível obter taxas de acerto em 100% nas etapas de treinamento e testes para os diferentes tipos de cargas do motor.

Conforme visto anteriormente na Figura 5, os descritores baseados em dados estatísticos apresentam boa separabilidade interclasses e baixa dispersão dos dados intraclasses. Os bons resultados de classificação nas etapas de treinamento e testes comprovam que não houve um *sobreajuste* (*overfitting*) dos modelos utilizados.

6. CONCLUSÕES

Nesta pesquisa, são utilizadas duas metodologias para extração de características através de sinais de vibração

Tabela 8. Resultados com classificador DT

Potência[hp]	Parâmetro	Acurácia [%]	
		Treino	Teste
0	Entropy	100,00	100,00
	Gini	100,00	100,00
1	Entropy	99,74	100,00
	Gini	99,74	99,42
2	Entropy	100,00	100,00
	Gini	100,00	100,00
3	Entropy	100,00	100,00
	Gini	100,00	100,00

Tabela 9. Resultados com classificador k-NN

Potência[hp]	Parâmetro	Acurácia [%]	
		Treino	Teste
0	k=1	92,18	91,30
	k=3	92,52	93,70
	k=5	93,54	92,10
1	k=1	100,00	100,00
	k=3	100,00	100,00
	k=5	100,00	100,00
2	k=1	100,00	100,00
	k=3	100,00	100,00
	k=5	100,00	100,00
3	k=1	99,74	100,00
	k=3	100,00	100,00
	k=5	99,74	99,74

Tabela 10. Resultados com classificador SVM

Potência[hp]	Parâmetro	Acurácia [%]	
		Treino	Teste
0	Gaussian	93,20	92,90
	linear	86,05	86,50
	Polynomial	98,98	100,00
1	Gaussian	100,00	100,00
	linear	99,74	100,00
2	Polynomial	100,00	100,00
	Gaussian	100,00	100,00
3	linear	99,47	100,00
	Polynomial	100,00	100,00
3	Gaussian	99,74	100,00
	linear	97,88	98,80
	Polynomial	99,74	100,00

Tabela 11. Resultados com classificador RF

Potência[hp]	Parâmetro	Acurácia [%]	
		Treino	Teste
0	5 trees	99,90	100,00
	10 trees	99,98	100,00
	15 trees	99,99	100,00
	20 trees	100,00	100,00
1	5 trees	99,94	99,40
	10 trees	100,00	100,00
	15 trees	100,00	100,00
	20 trees	100,00	100,00
2	5 trees	99,94	100,00
	10 trees	100,00	100,00
	15 trees	100,00	100,00
	20 trees	100,00	100,00
3	5 trees	99,94	100,00
	10 trees	100,00	100,00
	15 trees	100,00	100,00
	20 trees	100,00	100,00

num simulador de equipamentos dinâmicos, o qual possui diferentes naturezas de operação como condição normal (sem falha) e diferentes falhas em rolamentos, as quais são

comuns em equipamentos rotativos utilizados em instalações industriais. A partir dos resultados experimentais obtidos, foi possível comprovar que a análise dos sinais originais com dimensão reduzida com a utilização da PCA não permite a obtenção de bons modelos de classificação, tendo em vista o arranjo complexo das amostras no espaço amostral, onde apresentam pouca separabilidade. Outro fator que cooperou para os resultados ruins com os descritores da PCA foi a perda de informação relevante para a modelagem quando da redução das dimensões do espaço amostral. O melhor resultado utilizando a PCA (88,90%) não garante a confiabilidade e a precisão de um diagnóstico inteligente. Os descritores baseados em dados estatísticos, em geral, obtiveram bons resultados de classificação (100%), sendo que os melhores resultados foram obtidos com o classificador *DT*, utilizando *entropy* como regra de decisão, e com o classificador *RF* com uma coleção de dez árvores. O bom resultado apresentado com os descritores baseados em dados estatísticos comprovam a boa separabilidade dos dados, o que facilita o desempenho dos algoritmos de classificação utilizados. Os resultados dos testes mostram a viabilidade de se implementar sistemas de diagnósticos precisos e confiáveis no monitoramento de equipamentos dinâmicos utilizando técnicas de *machine learning* para o reconhecimento de padrões de falhas.

REFERÊNCIAS

- Agrawal, P. and Jayaswal, P. (2019). Diagnosis and classifications of bearing faults using artificial neural network and support vector machine. *Journal of The Institution of Engineers (India): Series C*, 1–12. doi:10.1007/s40032-019-00519-9.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. doi:10.1023/A:1010933404324. URL <https://doi.org/10.1023/A:1010933404324>.
- Chen, Z., Gryllias, K., and Li, W. (2019). Mechanical fault diagnosis using Convolutional Neural Networks and Extreme Learning Machine. *Mechanical Systems and Signal Processing*, 133, 106272. doi:10.1016/j.ymssp.2019.106272.
- Cortes, C. and Vapnik, V. (1995). Support-vector networks. In *Machine Learning*, 273–297.
- Dey, S., Roy, S.S., Samanta, K., Modak, S., and Chatterjee, S. (2019). Autocorrelation based feature extraction for bearing fault detection in induction motors. In *2019 International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, 1–5.
- Dougherty, G. (2013). *Pattern Recognition and Classification: An Introduction*. Springer Publishing Company, Incorporated.
- Duda, R.O., Hart, P.E., and Stork, D.G. (2001). *Pattern Classification*. Wiley, New York, 2 edition.
- Nayana, B.R. and Geethanjali, P. (2017). Analysis of statistical time-domain features effectiveness in identification of bearing faults from vibration signal. *IEEE Sensors Journal*, 17(17), 5618–5625.
- Rauber, T.W., de Assis Boldt, F., and Varejão, F.M. (2015). Heterogeneous feature models and feature selection applied to bearing fault diagnosis. *IEEE Transactions on Industrial Electronics*, 62(1), 637–646.
- Tahir, M.M., Khan, A.Q., Iqbal, N., Hussain, A., and Badshah, S. (2017). Enhancing fault classification accuracy of ball bearing using central tendency based time domain features. *IEEE Access*, 5, 72–83.
- Toma, R., Prosvirin, A., and Kim, J. (2020). Bearing fault diagnosis of induction motors using a genetic algorithm and machine learning classifiers. *Sensors*. doi:10.3390/s20071884.
- Webb, A. (2003). *Statistical Pattern Recognition*. Wiley InterScience electronic collection. Wiley.
- Zhang, R., Peng, Z., Wu, L., Yao, B., and Guan, Y. (2017). Fault diagnosis from raw sensor data using deep neural networks considering temporal coherence. *Sensors*, 17, 549. doi:10.3390/s17030549.
- Zhang, S., Zhang, S., Wang, B., and Habetler, T.G. (2019). Machine learning and deep learning algorithms for bearing fault diagnostics - A comprehensive review. *CoRR*, abs/1901.08247. URL <http://arxiv.org/abs/1901.08247>.
- Zhou, Q., Shen, H., Zhao, J., Liu, X., and Xiong, X. (2019). Degradation state recognition of rolling bearing based on k-means and cnn algorithm. *Shock and Vibration*, 2019, 1–9. doi:10.1155/2019/8471732.