

Machine Learning Algorithms for Remaining Useful Life Prediction of Rolling Bearings

Débora Zumpichiatti¹, Janaína Gomide¹

¹Instituto Politécnico - Universidade Federal do Rio de Janeiro (UFRJ)
Av. Aluizio da Silva Gomes, 50 – Macaé – RJ – Brasil

<deborazumpichiattil6@gmail.com>, <janainagomide@gmail.com>

Abstract. *The increasing complexity of mechanical systems changes the methods used to monitor and analyze how these systems age. The goal of this paper is to predict of the remaining useful life of equipments using a data-driven prognostic approach and machine learning algorithms. The dataset used presents temperature and vibration data from tests to run-to-failure of bearings. The proposed methodology was evaluated and the importance of a robust data processing phase was verified. The results obtained for data sets judged as appropriate by the methodology presented similar or superior results to the related works.*

Resumo. *O aumento da complexidade dos sistemas mecânicos muda drasticamente os métodos usados para monitorar e analisar como esses sistemas envelhecem. O objetivo desse trabalho é realizar a previsão do tempo de vida útil restante de equipamentos utilizando uma abordagem de prognóstico baseada em dados e algoritmos de aprendizado de máquina. O conjunto de dados utilizado apresenta dados de temperatura e vibração de testes até a falha de rolamentos. A metodologia proposta foi avaliada e constatou-se a importância de uma fase de tratamento de dados robusta. Os resultados obtidos para conjuntos de dados julgados como apropriados pela metodologia apresentaram resultados similares ou superiores aos trabalhos relacionados.*

1. Introdução

A complexidade das máquinas e indústrias aumentou ao longo dos últimos anos, impactando diretamente na quantidade e capacidade de monitorar as variáveis desses sistemas. Sob esse novo panorama surge o modelo de gestão chamado de Prognóstico e Gestão da Saúde (*Prognostic and Health Management - PHM*) que propõe monitorar o equipamento e conhecer a maneira como o mesmo envelhece, não apenas executando manutenções corretivas à medida que o equipamento falha. Segundo [Bougacha et al. 2020], a execução do modelo PHM de manutenção baseia-se em três fases principais: (1) Aquisição dos dados e avaliação de saúde, realizada através do uso de sensores que monitoram diferentes variáveis do sistema gerando um conjunto de dados passíveis de serem analisados; (2) Diagnóstico e prognóstico; (3) Tomada de decisão.

O propósito do prognóstico é estimar a vida útil remanescente (*Remaining Useful Life - RUL*) do sistema, de maneira a possibilitar a realização das devidas manutenções e evitar falhas em proporções catastróficas. Em [Patil et al. 2018] os autores relatam que o prognóstico pode ser dividido em duas abordagens, sendo a primeira baseada em modelo (*model-based*) e a segunda baseada em dados (*data-driven*). Em

[Sloukia et al. 2013b] os autores definem ainda uma terceira abordagem baseada na experiência (*experience-based*).

A abordagem baseada em modelo está diretamente ligada a aplicação de modelos analíticos para a previsão de RUL que, apesar de fornecer resultados mais precisos, torna-se de difícil aplicação visto que sistemas reais são não lineares e os mecanismos de degradação são estocásticos. A abordagem baseada na experiência usa os dados de feedback reunidos durante um período de tempo significativo, dependendo portanto da existência de uma quantidade massiva de dados. Por fim, a abordagem baseada em dados objetiva trabalhar com a identificação de padrões nos dados brutos obtidos na fase de aquisição de dados, de modo a obter informações relevantes sobre o sistema.

A metodologia PHM promete benefícios significativos ao reduzir o custo e o tempo de operação de manutenção. Esses benefícios estão relacionados à tomada de decisão baseada em informações prognósticas bem estruturadas por meio de diferentes metodologias [Iyer et al. 2006]. Esse artigo propõe uma metodologia de prognóstico baseada em dados que objetiva prever o RUL de rolamentos por meio da aplicação de algoritmos de aprendizado de máquina. Os experimentos foram realizados com o conjunto de dados produzido pela *IEEE Reliability Society* e *FEMTO-ST Institute (Franche-Comté Electronics Mechanics Thermal Science and Optics – Sciences and Technologies)* para o *IEEE PHM 2012 Data Challenge*. A metodologia é composta por oito etapas: estruturação dos dados, extração das características, ranking das características, filtro de média móvel, análise dos componentes principais, treinamentos de modelos, previsões de tempo de vida útil e avaliação dos resultados com base nas métricas selecionadas.

2. Trabalhos relacionados

Em 2020, [Bougacha et al. 2020] propôs uma revisão bibliográfica através da qual pode-se observar uma síntese dos trabalhos voltados ao PHM publicados no intervalo de tempo entre 2006 e 2018. A evolução das publicações mostra o crescente interesse em buscar novas metodologias de pesquisa voltadas para a aplicação do PHM como suporte na tomada de decisão em diversas indústrias. Uma das aplicações notáveis citada foi implementada por [Lin et al. 2016] no planejamento de manutenção de ferrovias, onde usou-se um processo de decisão de Markov (*Markov Decision Process - MDP*) parcialmente observável com o objetivo de planejar as intervenções de manutenção da fonte de alimentação de tração, com base em seu RUL.

Uma proposta de melhoria da utilização do Deep Learning como método de classificação fractográfica foi feita por [Bastidas-Rodríguez et al. 2019]. Utilizando quatro conjuntos de dados composto por imagens que correspondiam a quatro modos de falha mecânica (falha dúctil, frágil, por fadiga e fadiga por corrosão) e integrando a extração de características artesanais (*handcrafted features*) a aplicação de redes neurais profundas o modelo proposto mostrou desempenho satisfatório ao realizar a classificação de texturas. Os resultados alcançados trazem um novo fator a ser levado em consideração na hora da tomada de decisão do especialista que, de maneira geral, utilizaria apenas marcadores de textura e superfície para determinar o tipo de fratura.

A busca por reduzir cada vez mais o tempo de parada e número de falhas de equipamentos é uma constante no que tem sido conhecido como a Indústria 4.0, onde a manutenção preditiva tem estado em foco. Esse pensamento se reflete em uma série

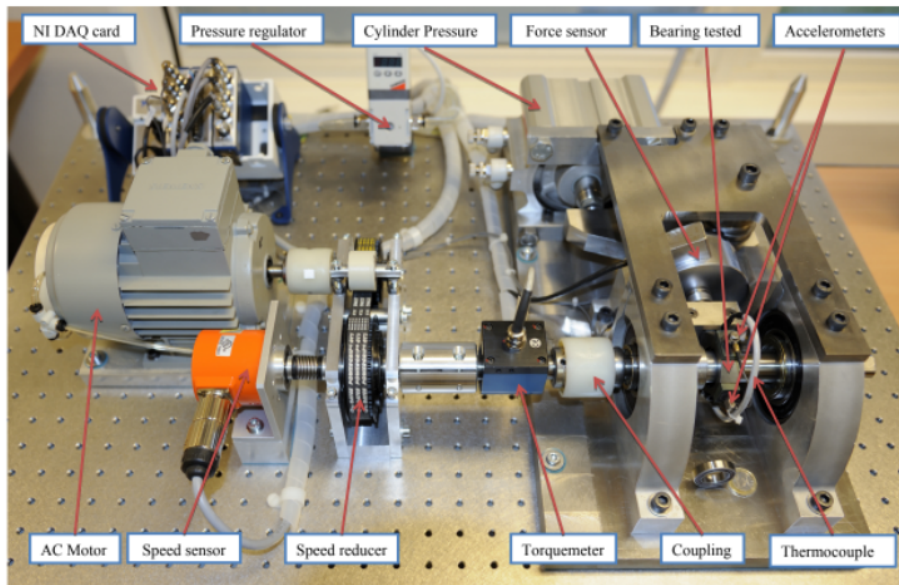


Figura 1. PRONOSTIA

de publicações que visam aplicar algoritmos cada vez mais robustos associados a sistemas complexos de aquisição de dados na monitoração de elementos críticos. Em [Lee et al. 2019] foi proposto um modelo aplicado a duas máquinas nas quais foram selecionados elementos críticos, os mesmos foram monitorados e avaliados com base em métricas pré-determinadas (normal, atenção e falha) e, por fim, foram aplicados algoritmos de aprendizado de máquina para classificação do status da máquina. De maneira similar, os trabalhos apresentados em [Sutrisno et al. 2012] e [Sloukia et al. 2013b] propuseram metodologias de previsão do RUL de rolamentos, dedicados a caracterização do estado de degradação e saúde desses elementos. São diversas as abordagens e aplicações e os resultados obtidos ratificam a importância e relevância desse tema de pesquisa.

3. Conjunto de dados

O conjunto de dados utilizado nesse trabalho foi gerado pelo Instituto FEMTO-ST¹ e é descrito em detalhes no artigo [Nectoux et al. 2012]. Esses dados foram utilizados no *IEEE PHM 2012 Data Challenge* com o objetivo de estimar o RUL de rolamentos. Esse problema é relevante visto que está diretamente relacionado à maioria das falhas de máquinas rotativas, afetando de maneira significativa a disponibilidade, segurança e custo-benefício de sistemas e equipamentos mecânicos em diversas indústrias.

Os experimentos para geração dos dados foram realizados em uma plataforma experimental de laboratório, o PRONOSTIA, Figura 1, onde foram simuladas situações reais de degradação de rolamentos até sua falha total em algumas horas. O sistema de testes é composto por três partes principais: uma parte rotativa, uma parte de geração de degradação (com uma força radial aplicada no rolamento testado) e uma parte de aquisição de dados. Obteve-se duas medidas de caracterização da degradação dos rolamentos: a vibração e a temperatura. Os sensores de vibração consistem em dois acelerômetros em miniatura posicionados a 90° um do outro; o primeiro é colocado no eixo vertical e o

¹FEMTO-ST Institute: <http://www.femto-st.fr/>

segundo é colocado no eixo horizontal. Já o sensor de temperatura é o chamado RTD (*Resistance Temperature Detector*) que é colocado dentro de um furo próximo ao anel do rolamento externo do sistema.

Foram consideradas 3 condições de operação para a parte rotativa e de degradação do sistema: Condição de operação 1: 1800 rotações por minuto (rpm) e 4000 Newton (N); Condição de operação 2: 1650 rpm e 4200 N; Condição de operação 3: 1500 rpm e 5000 N. A aquisição dos dados foi realizada de acordo com as seguintes características:

- Sinais de vibração (horizontal e vertical)
 - Frequência de amostragem: 25,6 kHz
 - Gravações: 2560 amostras (1/10 s) gravadas a cada 10 segundos
- Sinais de temperatura
 - Frequência de amostragem: 10 Hz
 - Gravações: 600 amostras gravadas a cada minuto

Foram realizados inicialmente 6 testes até a falha que compõem o conjunto de treinamento (*Learning Sets*) e, posteriormente, mais 11 testes que compõem o conjunto de teste (*Test Set*), Tabela 1. Esses conjuntos de dados foram definidos arbitrariamente pelo instituto responsável pela geração dos dados, bem como a quantidade de conjuntos para treinamento e teste e estão disponíveis online².

Tabela 1. Divisão e nomenclatura do conjunto de dados.

Dados	Condições de operação		
	Condição de operação 1	Condição de operação 2	Condição de operação 3
Conjuntos de Treinamento	Rolamento1.1	Rolamento2.1	Rolamento3.1
	Rolamento1.2	Rolamento2.2	Rolamento3.2
Conjuntos de Teste	Rolamento1.3	Rolamento2.3	Rolamento3.3
	Rolamento1.4	Rolamento2.4	
	Rolamento1.5	Rolamento2.5	
	Rolamento1.6	Rolamento2.6	
	Rolamento1.7	Rolamento2.7	

4. Metodologia

A metodologia proposta nesse trabalho é composta por seis etapas, conforme Figura 2 é descrita em detalhe a seguir.

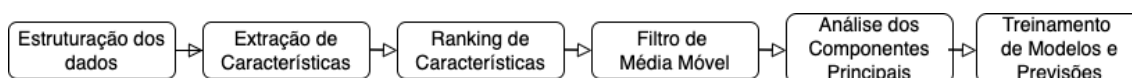


Figura 2. Metodologia proposta para a previsão de RUL de rolamentos.

²Conjunto de dados disponível em: <https://www.kaggle.com/datasets/alanhabrony/ieee-phm-2012-data-challenge>

4.1. Estruturação dos dados

Os conjuntos de dados gerados pelo Instituto FEMTO-ST foram disponibilizados em arquivos ASCII. A primeira fase da metodologia consistiu em organizar esses conjuntos utilizando uma única unidade representativa para o tempo e transformando as diversas colunas em apenas uma constando o marco temporal. Isso foi necessário devido as diferentes características de aquisição de dados para o caso da vibração e de temperatura. Enquanto os sinais de temperatura eram registrados a cada décimo de segundo, os sinais de vibração eram registrados a cada microssegundo.

O próximo passo foi determinar o RUL dos dados de treinamento. Esse valor foi estimado, para cada marco de tempo, como a diferença entre o tempo atual e o marco de tempo registrado na iminência da falha. Dessa forma pode-se determinar o vetor RUL composto pelo tempo de vida útil registrado durante o experimento e que posteriormente será utilizado como parâmetro de aprendizado para os algoritmos de aprendizado de máquina. Para ilustrar um exemplo do comportamento da vibração e da temperatura ao longo do tempo foi utilizado o primeiro conjunto de teste (Rolamento1.1) para gerar gráficos. As Figuras 3(a) e 3(b) apresentam a evolução temporal dos dados de vibração nos eixos x e y do primeiro teste realizado. A Figura 3(c) apresenta a evolução da temperatura para os mesmos dados. Pode-se observar que, diferente dos sinais de vibração, os dados de temperatura registrados possuíam pouca variação temporal. Tal comportamento mostrou que os dados de vibração apresentavam informação de maior qualidade no que dizia respeito a saúde do equipamento em análise, sendo escolhido para a aplicação.

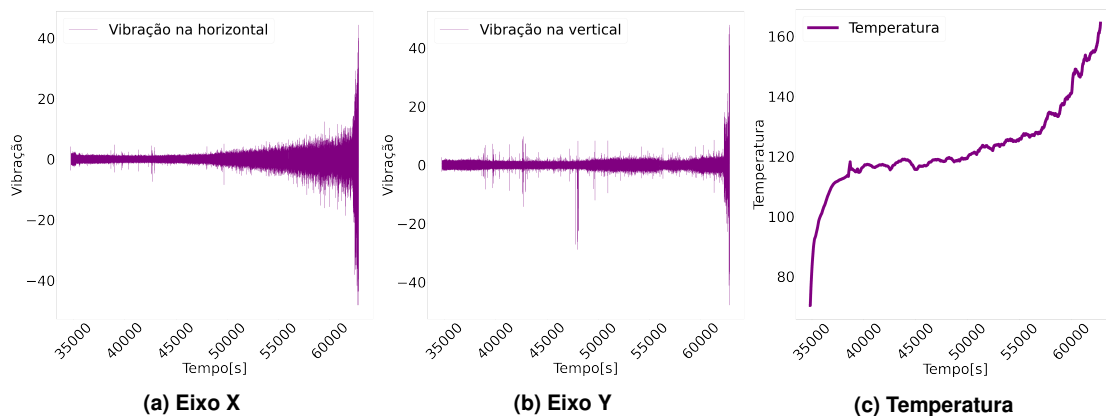


Figura 3. Vibração nos eixos X e Y e temperatura do Rolamento1.1.

4.2. Extração das Características

A fase de extração das características consiste em extrair dos dados brutos informações representativas sobre a saúde do sistema. Essas novas características podem ser divididas em 3 classes principais, segundo [Sloukia et al. 2013b] : características no domínio do tempo (*Time Domain*), características no domínio da frequência (*Frequency Domain*) e uma mistura dessas abordagens (*Time-Frequency Domain*).

Nesse trabalho foram extraídas características no domínio do tempo relacionadas ao sinal de vibração. Os sinais de vibração utilizados são a vibração na horizontal (eixo X) e vibração na vertical (eixo Y). Para obter as características ao longo do tempo foi utilizada uma janela de 2560 amostras, que equivale a quantidade de dados adquiridos

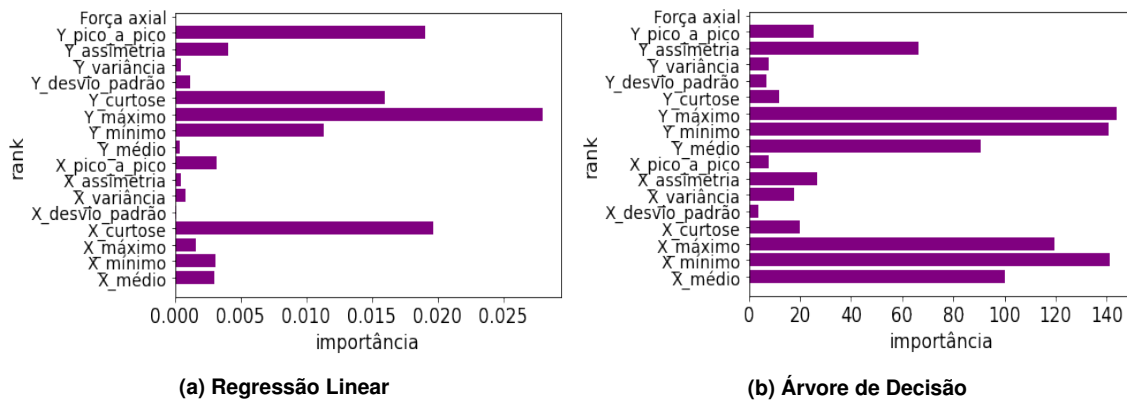


Figura 4. Ranking das características

em um intervalo de 10 segundos. Foram geradas oito características: média, mínimo, máximo, curtose, desvio padrão, variância, assimetria e valor de pico a pico para cada um dos sinais de vibração. Dessa forma, o conjunto de dados usado passa a ser composto por dezesseis colunas contendo as características no domínio do tempo, duas colunas compostas pelas condições de operação e uma coluna com o RUL estimado.

4.3. Ranking das Características

A fim de entender a representatividade das informações fornecidas por cada característica extraída do conjunto de dados foi gerado um ranking das características. O ranking é calculado baseado em um *score* que representa a importância da característica, quanto maior esse valor mais efeito tem no modelo de aprendizado de máquina.

O *score* de cada característica foi calculado utilizando o algoritmo de Regressão Linear e de Árvore de Decisão. A Regressão Linear ajusta os dados a um modelo linear, reduzindo a soma residual dos quadrados entre os alvos observados no conjunto de dados e os alvos previstos pela aproximação linear, [Géron 2019]. O algoritmo fornece o coeficiente da regressão que permite isolar o papel de cada variável das demais no modelo. Esse coeficiente representa a média da mudança na variável resposta para uma unidade de mudança na variável preditora, mantendo as outras preditoras na constante do modelo. Dessa maneira, podemos conhecer a importância de cada característica dentro da Regressão Linear. O algoritmo de Árvore de Decisão é uma estrutura de dados recursiva que define um modelo baseado em árvore com um nó folha que corresponde a um valor ou classe e um nó de decisão que contém um teste sobre algum atributo, [Zaki et al. 2014]. O algoritmo da Árvore de Decisão fornece um coeficiente de importância, que permite gerar um ranking com a importância das características.

As Figuras 4(a) e 4(b) mostram o ranking das características criado com a Regressão Linear e com a Árvore de Decisão, respectivamente. As características foram nomeadas, na Figura 4, seguindo o padrão eixo (X ou Y) seguido o nome da medida estatística. Além disso, a característica *Força axial* representa a carga a que o rolamento estava submetido durante o teste. Apesar do comportamento observado em cada algoritmo ser diferente, pode-se observar que a vibração máxima e mínima apresenta um alto nível de importância. Além disso, o Y de pico a pico foi a segunda característica com importância mais marcante para os dois algoritmos. Esse comportamento faz sentido visto

que está sendo analisado o comportamento vibracional da falha que deve ser marcado por altas amplitudes na iminência de sua ocorrência.

4.4. Filtro de Média Móvel

A imensa variabilidade dos dados vibracionais torna a identificação de padrões nos dados brutos uma tarefa complexa. Podemos observar um aumento na amplitude do movimento de vibração com o passar do tempo. Esse comportamento fica mais marcante na iminência da falha e pode ser observado na Figura 3. A fim de tornar a tendência desses dados dispostos ao longo do tempo mais evidente e facilitar o processo de identificação de padrões de falha no sistema aplicou-se a média móvel ao conjunto de dados composto pelas novas características extraídas dos dados brutos.

Como o próprio nome sugere, a média móvel tira uma média dos valores em um dado intervalo de tempo. Para essa aplicação foi selecionado trabalhar com intervalos de 6 segundos. Após a aplicação do filtro de média móvel é possível definir quais são os conjuntos de dados de maior qualidade para a aplicação da metodologia. Os conjuntos de dados que seguiam um padrão de falha mais claro, com aumento gradativo da amplitude de vibração ao longo do tempo, sem maiores variabilidade, foram definidos como os dados de maior qualidade.

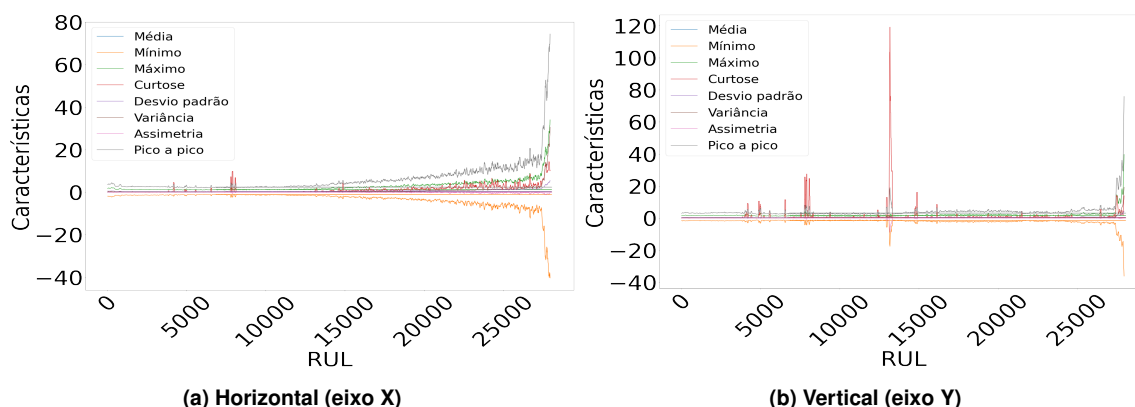


Figura 5. Características ao longo do tempo para o conjunto de dados Rolamento1.1

A Figura 5 apresenta o resultado após essa aplicação, também para o primeiro teste. Nesse caso é possível observar uma alteração nos valores do eixo x se comparado a figuras anteriores. Isso ocorre após a aplicação do filtro de média móvel visto que o mesmo agrupo os dados em grupos menores e aplica a média aritmética a esses. Aplicar o filtro de média móvel nesse caso significa ter uma redução na quantidade de dados para 1/6 de seu valor original.

4.5. Análise dos Componentes Principais

Problemas reais de Aprendizado de Máquina envolvem uma grande quantidade de características para cada instância de treinamento. Isso se torna um problema a medida que o treinamento converte-se em uma atividade extremamente lenta, além de ser mais difícil de encontrar uma boa solução, [Géron 2019]. As técnicas de redução de dimensionalidade surgem como a solução desse problema, transformando um problema intratável

em um tratável. Ainda segundo [Géron 2019], a Análise de Componentes Principais (*Principal Components Analysis - PCA*) é um dos métodos mais populares aplicados na redução de dimensionalidade por ser uma técnica linear e não paramétrica. Baseado na variância dos dados ele cria uma nova representação dos dados com menor dimensão, mantendo a variância e representatividade. Para a aplicação foi escolhido testar diferentes números de componentes, a fim de observar o comportamento dos dados e escolher a melhor representação.

4.6. Treinamento de modelos e Previsões

Na fase de treinamento de modelos foram selecionados algoritmos de aprendizado de máquina aplicados para regressão. Segundo trabalho apresentado em [Sloukia et al. 2013a] os modelos mais aplicados para metodologias voltadas para prognóstico são as Redes Neurais, Máquina de Vetores de Suporte e a Teoria *Fuzzy*. Em [Patil et al. 2018] os autores citam o bom desempenho de modelos *ensemble* para aplicações similares.

Dessa forma, nesse trabalho foram selecionados e treinados três tipos de modelos: o Máquina de Vetores de Suporte para Regressão (SVR) é um algoritmo de aprendizado supervisionado usado para prever valores discretos que tem como ideia básica encontrar o hiperplano que possui o número máximo de pontos; o *Gradient Boosting* (GB) e Florestas Aleatórias (RF) são algoritmos que utilizam o método *ensemble*, técnica que combina o resultado de múltiplos modelos em busca de produzir um melhor modelo preditivo. Enquanto as Florestas Aleatórias combinam a saída de várias árvores de decisão para chegar a um único resultado o *Gradient Boosting* cria árvores decisões com o objetivo de prever o valor dos erros da árvore anterior, e utiliza o valor do erro previsto na definição de seu resultado final.

O conjunto de dados utilizado já está separado em treino e teste, conforme descrito na seção 3. Para treinar os modelos foram utilizados os dados de treino e os modelos foram avaliados utilizando os dados de teste. Os dados de teste passaram por todas as fases de tratamento propostas ao decorrer da metodologia. De posse dos algoritmos treinados e conjuntos de dados de teste tratados foram realizadas as previsões e avaliação dos resultados que serão apresentados a seguir.

5. Resultados

A metodologia proposta fez o tratamento do conjunto de dados e alterou os dados de treinamento em diversas fases. Inicialmente, os dados selecionados a serem usados eram compostos por marcos temporais de hora, minuto, segundo e micro segundo, além dos dados vibracionais extraídos do setup na horizontal e na vertical. Após a fase de extração das características passou-se para uma nova configuração dos dados, composta por oito características no domínio do tempo extraídas dos dados da horizontal, oito características no domínio do tempo extraídas dos dados da vertical, a RUL estimada para cada marco temporal, a rotação em RPM para a dada configuração do sistema e a carga aplicada. Após a aplicação do filtro de média móvel não observou-se aumento ou diminuição na quantidade de características do sistema, no entanto alcançou-se a diminuição da quantidade de dados, ou seja, da quantidade de instâncias a serem usadas no treinamento de modelos. Foi selecionado trabalhar com uma janela rolante (*rolling window*) de 6 dados e, por consequência, houve uma redução do conjunto de dados a 1/6 de seu valor original.

Tabela 2. Modelos de Florestas aleatórias treinados.

PCA	RF
número de componentes=2,4	Padrão Máxima profundidade=10 e número de árvores=10,20,50,100 Máxima profundidade=20 e número de árvores=10,20,50,100 Máxima profundidade=100 e número de árvores=10,20,50,100

Na fase do Análise dos Principais Componentes (PCA) foram testadas duas opções: dois ou quatro componentes. Ao aplicar-se a primeira redução, de dezoito para quatro dimensões, observou-se que os dois primeiros componentes representavam, em média, um percentual de 94 do total dos dados, indicando que redução de dezoito para duas dimensões poderia ser satisfatória.

Foram realizadas diferentes configurações dos modelos de aprendizado de máquina conforme a variação dos hiper-parâmetros utilizados. Para o PCA foi variado o número de componentes utilizando os valores 2 e 4. Os hiper-parâmetros utilizados no algoritmo de Florestas Aleatórias foram de profundidade máxima e número de árvores, conforme apresentados na Tabela 2. Para o algoritmo de *Gradient Boosting*, foram utilizados os parâmetros padrão da biblioteca *Scikit-Learn* sendo a função de perda o erro quadrático, número de estimadores igual a 100 e máximo de profundidade igual a 3. Para o algoritmo SVR foi utilizado o Kernel da função de base radial (RBF). No total foram geradas 30 configurações diferentes de modelos.

Os dados de treinamento foram reorganizados a depender das condições de operação citadas na Tabela 1 e, portanto, foi usado para treinamento três grandes conjuntos de dados, chamados de configuração 1, configuração 2 e configuração 3. Cada um desses conjuntos de dados de treinamento foi usado para treinar todos os 30 modelos treinados, dos quais dois são o *Gradiente Boosting* com parâmetros padrão e 2 e 4 componentes no PCA, dois são o Máquina de Vetores de Suporte para regressão com parâmetros padrão a 2 e 4 componentes do PCA. As outras variações feitas com o as Florestas Aleatórias estão detalhadas na Tabela 2.

Para avaliar o desempenho de cada modelo treinado foi necessário processar todos os dados dos onze conjuntos de teste e realizar a estimativa de previsão de RUL para cada um deles. Aplicando-se os 30 modelos obteve-se 330 estimativas de RUL. Para avaliar o desempenho de cada modelo aplicou-se o erro percentual para cada experimento (E_i) conforme sugerido pelo *IEEE PHM 2012 Prognostic Challenge*.

$$E_i = 100 \left(\frac{ActRUL_i - RUL_i}{ActRUL_i} \right)$$

onde, E_i é o erro percentual no experimento i , RUL_i é o valor de RUL estimado para o experimento i , $ActRUL_i$ é o valor real de RUL. Portanto, são consideradas como melhores previsões aquelas cujo erros percentuais são maiores ou iguais a zero, ou seja, previsões que caracterizam antecipação da falha. Além disso possuir o módulo do erro percentual mais próximo de zero é o ideal visto que, quanto menor o erro percentual, mais a previsão se assemelha ao valor real esperado.

A Tabela 3 apresenta os erros percentuais obtidos usando os parâmetros padrão,

em negrito estão as previsões com melhores desempenhos. Os resultados obtidos pelas Florestas Aleatórias cujos parâmetros foram alterados não apresentaram um padrão de melhoria se comparado aos parâmetros padrão. Os melhores resultados foram obtidos para os conjuntos de dados definidos como os dados de maior qualidade, conforme citado da seção 4.4, e são eles o Rolamento1.3, Rolamento1.7 e Rolamento2.3. Observou-se também que o desempenho dos modelos treinados com quatro componentes do PCA obtiveram melhores desempenhos ao serem submetidos a métrica selecionada.

Tabela 3. Erro percentual calculado.

	Quantidade de componentes no PCA					
	2			4		
	SVR	GB	RF	SVR	GB	RF
Rolamento1.3	-55,174	12,175	-8,569	-55,090	-1,100	-9,74
Rolamento1.4	-2511,648	-1217,756	-867,109	-2511,597	-755,315	-866,14
Rolamento1.5	-487,776	-348,074	-417,944	-488,059	-508,900	-423,09
Rolamento1.6	-554,480	-391,700	-596,199	-554,823	-546,817	-598,89
Rolamento1.7	-24,223	14,420	-8,791	-24,279	-7,362	-9,36
Rolamento2.3	43,581	-30,777	10,242	43,638	-13,368	9,74
Rolamento2.4	-196,421	-202,588	-48,770	-196,274	-32,880	-51,73
Rolamento2.5	-38,979	49,978	78,026	-38,534	63,520	77,96
Rolamento2.6	-227,745	-56,438	-23,388	-226,949	-9,474	-25,33
Rolamento2.7	-690,683	-1133,500	-939,414	-687,075	-974,210	-979,38
Rolamento3.3	-574,608	59,041	-192,457	-575,770	-363,836	-70,58

A Tabela 4 apresenta resultados obtidos pelos trabalhos relacionados apresentados em [Sutrisno et al. 2012] e [Patil et al. 2018]. Neles foram aplicados, respectivamente, uma metodologia de detecção de anomalia na frequência e dois métodos ensemble, as Florestas aleatórias e o *Gradient Boosting*. Pode-se observar que os melhores resultados obtidos pela metodologia proposta nesse trabalho se igualam, ou até superam, resultados obtidos por esses trabalhos relacionados (Rolamento1.3, Rolamento1.7, Rolamento2.3, Rolamento2.5). No entanto, metodologias mais complexas, tiveram melhores desempenhos nos casos de dados com maior variabilidade. Em [Patil et al. 2018] a melhoria de desempenho pode ser justificada pela aplicação de uma técnica de eliminação recursiva de características, usando validação cruzada, e criação de novos conjuntos de dados para o treinamento dos algoritmos, diferenciando-o da metodologia proposta nesse trabalho. Através da análise dos resultados apresentados nas Tabelas 3 e 4 observa-se que o *Gradient Boosting* foi o algoritmo que apresentou melhor desempenho para esse tipo de aplicação.

6. Conclusões

Nesse trabalho foi apresentada uma metodologia proposta para realizar a previsão do tempo de vida útil restante de equipamentos mecânicos a partir de dados de vibração de testes realizados até a falha. Essa metodologia foi avaliada utilizando um conjunto de dados de rolamento. Pode-se perceber que a influência do ambiente de teste é muito forte em uma análise a partir de dados de vibração, além disso a repetibilidade das condições de teste podem não ser suficientes para obter resultados com mesmo erro percentual. Sendo

Tabela 4. Resultados obtidos por [Sutrisno et al. 2012] e [Patil et al. 2018].

	Detecção de anomalia na frequência [Sutrisno et al. 2012]	RF [Patil et al. 2018]	GB [Patil et al. 2018]
Rolamento1.3	37	-1,17	-0,61
Rolamento1.4	80	-2,46	2,34
Rolamento1.5	9	12,73	8,68
Rolamento1.6	-5	28,32	22,40
Rolamento1.7	-2	-4,90	-3,29
Rolamento2.3	64	29,47	19,86
Rolamento2.4	10	7,12	4,66
Rolamento2.5	-440	27,32	9,01
Rolamento2.6	49	3,89	0,41
Rolamento2.7	-317	43,21	49,22
Rolamento3.3	90	32,09	12,67

assim, mesmo para condições idênticas de teste os arquivos contavam com tempos até a falha diferentes e comportamentos vibracionais diferentes, dificultando a identificação de padrões. Os melhores desempenhos foram obtidos nos casos onde o comportamento dos dados segue a lógica esperada, onde com a evolução temporal e aproximação do momento da falha, observa-se maior amplitude de vibração, comportamento esse que marcava um padrão de falha. Os piores desempenhos no que diz respeito ao erro percentual da previsão se deram principalmente nos conjuntos de dados com maior variabilidade, onde houveram picos de vibração aleatórios ao longo de todo tempo, sem um padrão claro de falha.

A utilização de dois e quatro componentes na análise de componentes principais mostrou que não houve diferença marcante no modulo do erro percentual. Ao realizar uma análise mais profunda pode-se observar que os resultados obtidos para os quatro componentes obtiveram majoritariamente valores negativos que, para a aplicação, significa que a falha foi prevista de maneira tardia. Apesar disso tais resultados podem ser utilizados, na prática, para propor intervalos de tempo nos quais possam ocorrer a falha.

Pode-se ainda comparar os resultados obtidos com trabalhos relacionados que se propuseram a trabalhar com o mesmo conjunto de dados. Constatou-se a importância de uma fase de tratamento de dados robusta e que utilize diversas técnicas em conjunto, a fim de obter melhores resultados para dados com maior variabilidade. No entanto, os conjuntos de dados julgados como melhores pela metodologia proposta nesse trabalho obtiveram resultados similares ou até melhores que trabalhos relacionados.

Como trabalhos futuros, essa metodologia também pode ser aplicada a outros conjuntos de dados que contenham medições de outros tipos de equipamentos até a falha. Além disso, pode-se verificar o tratamento de dados mais complexo para dados julgados como não apropriados e a aplicação de outros algoritmos de aprendizado de máquina.

Referências

Bastidas-Rodríguez, M.-X., Prieto-Ortiz, F.-A., and Polanía, L. F. (2019). A textural deep neural network combined with handcrafted features for mechanical failure classifica-

- tion. In *2019 IEEE International Conference on Industrial Technology (ICIT)*, pages 847–852. IEEE.
- Bougacha, O., Varnier, C., and Zerhouni, N. (2020). A review of post-prognostics decision-making in prognostics and health management. *International Journal of Prognostics and Health Management*, 11(15):31.
- Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. "O'Reilly Media, Inc."
- Iyer, N., Goebel, K., and Bonissone, P. (2006). Framework for post-prognostic decision support. In *2006 IEEE Aerospace Conference*, pages 10–pp. IEEE.
- Lee, W. J., Wu, H., Yun, H., Kim, H., Jun, M. B., and Sutherland, J. W. (2019). Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data. *Procedia Cirp*, 80:506–511.
- Lin, S., Zhang, A., and Feng, D. (2016). Maintenance decision-making model based on pomdp for traction power supply equipment and its application. In *2016 Prognostics and System Health Management Conference (PHM-Chengdu)*, pages 1–6. IEEE.
- Nectoux, P., Gouriveau, R., Medjaher, K., Ramasso, E., Chebel-Morello, B., Zerhouni, N., and Varnier, C. (2012). Pronostia: An experimental platform for bearings accelerated degradation tests. In *IEEE International Conference on Prognostics and Health Management, PHM'12.*, pages 1–8. IEEE Catalog Number: CPF12PHM-CDR.
- Patil, S., Patil, A., Handikherkar, V., Desai, S., Phalle, V. M., and Kazi, F. S. (2018). Remaining useful life (rul) prediction of rolling element bearing using random forest and gradient boosting technique. In *ASME international mechanical engineering congress and exposition*, volume 52187. American Society of Mechanical Engineers.
- Sloukia, F., Bouarfa, R., Medromi, H., and Wahbi, M. (2013a). Bearings prognostic using mixture of gaussians hidden markov model and support vector machine. *International Journal of Network Security Its Applications*, 5:85–97.
- Sloukia, F., El Aroussi, M., Medromi, H., and Wahbi, M. (2013b). Bearings prognostic using mixture of gaussians hidden markov model and support vector machine. In *2013 ACS International Conference on Computer Systems and Applications (AICCSA)*, pages 1–4. IEEE.
- Sutrisno, E., Oh, H., Vasan, A. S. S., and Pecht, M. (2012). Estimation of remaining useful life of ball bearings using data driven methodologies. In *2012 IEEE conference on prognostics and health management*, pages 1–7. IEEE.
- Zaki, M. J., Meira Jr, W., and Meira, W. (2014). *Data mining and analysis: fundamental concepts and algorithms*. Cambridge University Press.