

Modelagem da Degradação em Turbinas de Aeronave sob Condições Reais de Voo

Joaquim O. F. Moura Filho* Marcelo E. da Silva**
Vandilberto P. Pinto*** Márcio A. B. Amora*

* Programa de Pós-Graduação em Engenharia Elétrica e Computação (PPGEEC), Campus Sobral, Universidade Federal do Ceará, CE, (e-mail: joaquim1905@alu.ufc.br, marcio@sobral.ufc.br)

** Engenharia Elétrica, Campus Sobral, Universidade Federal do Ceará, CE, (e-mail: marcelo.estevao@alu.ufc.br)

*** Instituto de Engenharias e Desenvolvimento Sustentável, Universidade da Integração Internacional da Lusofonia Afro-Brasileira (UNILAB), Redenção, CE, (e-mail: vandilberto@unilab.edu.br)

Abstract:

The paper performs a modeling of the degradation curves in aircraft turbines under real flight conditions and also a comparison between machine learning techniques based on decision trees. The algorithms used are: Decision Trees (DT), Random Forest (RF) and Gradient Boosting (GB). Coefficient of determination, mean square error and root mean square error are employed as performance evaluation methods. The presented results show the best performance of RF and GB in estimating the values. The coefficients of determination of the algorithms reached average values higher than 0.98, thus showing the efficiency of the proposed models to be used in this application.

Resumo: O trabalho realiza uma modelagem das curvas de degradação em turbinas de aeronaves sob condições reais de voo e também um comparativo entre técnicas de aprendizado de máquina baseadas em árvores de decisão. Os algoritmos utilizados são: Árvores de Decisão (AD), Floresta Aleatória (FA) e *Gradient Boosting* (GB). São empregados como métodos de avaliação de desempenho o coeficiente de determinação, o erro quadrático médio e a raiz do erro quadrático médio. Os resultados apresentados mostram o melhor desempenho da FA e do GB na estimação dos valores. Os coeficientes de determinação dos algoritmos atingiram valores médios superiores a 0,98, mostrando assim a eficiência dos modelos propostos para serem utilizados nesta aplicação.

Keywords: Machine learning; Decision trees; Aircraft turbine; Curve modeling.

Palavras-chaves: Aprendizagem de máquina; Árvores de decisão; Turbina de aeronave; Modelagem de curvas.

1. INTRODUÇÃO

A degradação de motores de aeronaves ou turbinas pode ocorrer a qualquer momento devido a uma variedade de problemas como corrosão e ranhuras nas lâminas da turbina. As degradações podem ocasionar em falhas nos motores das aeronaves ou turbinas que podem resultar em perdas econômicas significativas e até mesmo acidentes em casos extremos (Li et al., 2019).

A degradação pode ocorrer em qualquer categoria de sistema. Com o passar do tempo fatalmente um dispositivo ou equipamento para de funcionar. Sendo assim, torna-se necessário ao projetista construir modelos que representem a degradação do seu material ao longo do tempo. Isso torna seu produto mais confiável e sujeito a uma manutenção preditiva (Kang et al., 2020; Ameneh et al., 2017).

Os modelos de degradação são utilizados na criação de prognósticos, que possuem o intuito de monitorar a vida útil de um determinado equipamento. Também são em-

pregados nas manutenções baseadas em degradação, do inglês, *Degradation-Based Maintenance* (DBM), onde são empregados conceitos como manutenção perfeita, mínima e imperfeita (Kang et al., 2020).

A manutenção perfeita acontece quando além de reparar a falha em algum equipamento, são feitas análises na potencialidade da falha. Já a manutenção mínima ocorre quando é realizado apenas o reparo na parte defeituosa do equipamento, sem observar as causas que o levaram para aquele estado. Por fim, também existe o conceito de manutenção imperfeita. Neste caso a manutenção é feita para tornar o equipamento melhor do que as condições atuais deste (Lopes, 2012).

Nos trabalhos de Kang et al. (2020) e Ameneh et al. (2017) são feitas revisões e análises sobre abordagens utilizadas na realização de modelos de degradação. Em se tratando de sistemas aeronáuticos, que apresentam um modelo físico mais complexo, o uso de algoritmos que realizem a modelagem da degradação da falha ganharam destaque.

As turbinas são formadas por diversas partes e cada uma impacta na saúde da aeronave. Estes componentes foram feitos para terem uma longa durabilidade e confiabilidade, mas com o tempo de uso, geralmente ocorrem desgastes e fadigas que comprometem o seu desempenho (Zhang et al., 2020).

Em Saxena et al. (2008) é descrito como a propagação de danos pode ser modelada nos parâmetros de uma turbina a gás de aeronaves. Para fazer isso o autor emprega um modelo de simulação termodinâmica.

Já no artigo de Sogut et al. (2017), são analisados os efeitos da degradação em motores de aeronaves utilizando a *Exergy Analysis*. Com essa abordagem o autor trata dos efeitos das perdas de potência identificados durante o ciclo de vida de compressores e turbinas.

No trabalho de Zhang et al. (2020) é desenvolvido uma metodologia visando realizar o prognóstico do motor da aeronave. Nele é feito um processo de modelagem empregando o uso do método de análise de componentes (*Functional Principal Component Analysis - FPCA*).

Este artigo apresenta a utilização e comparação de técnicas de aprendizagem de máquina baseadas em árvores de decisão na criação de uma modelagem das curvas de degradação em turbinas de aeronaves.

Portanto, este trabalho possui a seguinte estrutura: a presente seção mostra uma breve introdução do tema proposto. Conceitos sobre o banco de dados utilizado são abordados na seção 2. Na seção 3 são apresentadas as técnicas que serão utilizadas para realizar a modelagem. Já na seção 4 são introduzidas as metodologias utilizadas para realizar o treinamento e o teste dos algoritmos. Na seção 5 são discutidos os resultados alcançados. Por fim, na sessão 6 é feita uma conclusão e são apresentadas as perspectivas para trabalhos futuros.

2. AQUISIÇÃO DE DADOS

O banco de dados utilizado neste trabalho foi retirado do repositório de prognóstico público da *National Aeronautics and Space Administration* (NASA) vindo do artigo de (Chao et al., 2021). Os dados foram criados de um modelo dinâmico de computador (*Commercial Modular Aero - Propulsion System Simulation - C-MAPSS*), no qual são empregados na simulação de degradação de motores comerciais a reação do tipo Turbofan. Em Frederick et al. (2007) é mostrado um guia com toda a modelagem usada no C-MAPSS.

O C-MAPSS é um modelo muito realista para a simulação de um motor Turbofan. Além de possuir o modelo termodinâmico do motor, o *software* também possui um modelo atmosférico e um sistema para gerenciamento de energia. Na Figura 1 retirado de (Chao et al., 2021) é apresentado o modelo.

O motor Turbofan, como representado na Figura 1, é composto por uma ventoinha (Fan), um compressor de baixa pressão (*Low-Pressure Compressor - LPC*), um compressor de alta pressão (*High-Pressure Compressor - HPC*), uma turbina de alta pressão (*High-Pressure Turbine - HPT*) e uma turbina de baixa pressão (*Low-Pressure Turbine - LPT*).

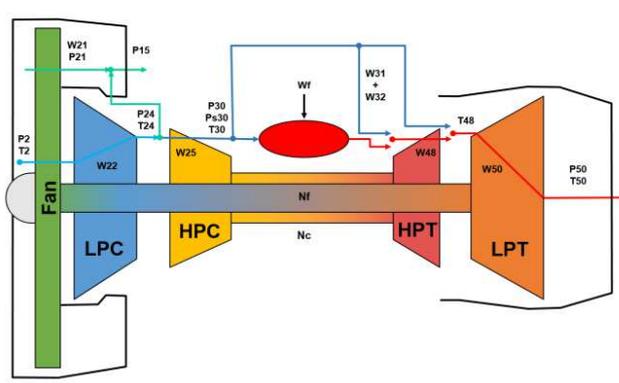


Figura 1. Representação do modelo C-MAPSS.

O modelo C-MAPSS é modelado por um sistema acoplado de equações não lineares, nas quais as entradas do modelo são divididas na descrição de condições operacionais w e nos parâmetros de saúde do modelo θ . As saídas são as estimativas das propriedades físicas \hat{x}_s e das propriedades não observadas \hat{x}_v que não fazem parte dos sinais de monitoramento de condição (Chao et al., 2021).

Por esse modelo ser hermético e com muitos parâmetros, se torna necessário o emprego de técnicas que possam manipular esses dados sem a necessidade da utilização do sistema complexo original, por essa razão, neste trabalho são utilizados métodos baseados em árvores de decisão. O sistema não linear pode ser escrito segundo Chao et al. (2021) como representado em (1):

$$[\hat{x}_s^{(t)}, \hat{x}_v^{(t)}] = S(w^{(t)}, \theta^{(t)}) \quad (1)$$

Nas Tabelas 1, 2, 3 e 4 são mostrados as descrições das variáveis empregadas no modelo C-MAPSS (Chao et al., 2021).

Tabela 1. Descrição da simbologia do modelo C-MAPSS para os parâmetros w .

Símbolo	Descrição	Unidade
alt	Altitude	ft
Mach	Número de Mach no voo	-
TRA	Ângulo do acelerador-resolvedor	%
T2	Temperatura total na entrada do ventilador	°R

Tabela 2. Descrição da simbologia do modelo C-MAPSS para os parâmetros θ .

Símbolo	Descrição	Unidade
fan_eff_mod	Modificador de eficiência do ventilador	-
fan_flow_mod	Modificador de fluxo do ventilador	-
LPC_eff_mod	Modificador de eficiência do LPC	-
LPC_flow_mod	Modificador de fluxo do LPC	-
LPC_flow_mod	Modificador de eficiência do HPC	-
HPC_flow_mod	Modificador de fluxo do HPC	-
HPT_eff_mod	Modificador de eficiência do HPT	-
HPT_flow_mod	Modificador de fluxo do HPT	-
LPT_eff_mod	Modificador de eficiência do LPT	-
LPT_flow_mod	Modificador de fluxo do LPT	-

O conjunto de dados fornece curvas sintéticas de degradação de nove motores Turbofan que possuem condições

iniciais de saúde desconhecidas. Foram colocados como entradas do modelo C-MAPSS as condições de voos reais registradas de um jato comercial (Chao et al., 2021).

Tabela 3. Descrição da simbologia do modelo C-MAPSS para os parâmetros \hat{x}_s .

Símbolo	Descrição	Unidade
Wf	Fluxo de combustível	pps
Nf	Velocidade física do ventilador	rms
Nc	Velocidade do núcleo físico	rms
T24	Temperatura total na saída LPC	°R
T30	Temperatura total na saída HPC	°R
T48	Temperatura total na saída HPT	°R
T50	Temperatura total na saída LPT	°R
P15	Pressão total no duto de desvio	psia
P21	Pressão total na saída do ventilador	psia
P24	Pressão total na saída LPC	psia
Ps30	Pressão estática na saída HPC	psia
P40	Pressão total na saída do queimador	psia
P50	Pressão total na saída LPT	psia

Tabela 4. Descrição da simbologia do modelo C-MAPSS para os parâmetros \hat{x}_v .

Símbolo	Descrição	Unidade
T40	Temperatura Total na saída do queimador	°R
P30	Pressão total na saída HPC	psia
P45	Pressão total na saída HPT	psia
W21	Fluxo do ventilador	pps
W22	Fluxo de LPC	lbm/s
W25	Fluxo para HPC	lbm/s
W31	HPT <i>coolant bleed</i>	lbm/s
W32	LPT <i>coolant bleed</i>	lbm/s
W48	Fluxo para fora do HPT	lbm/s
W50	Fluxo para fora do LPT	lbm/s
epr	Razão de pressão do motor (P50 / P2)	-
SmFan	Margem da parada do ventilador	-
SmLPC	Margem de estol LPC	-
SmHPC	Margem de perda de HPC	-
NRf	Velocidade corrigida do ventilador	rpm
NRc	Velocidade de núcleo corrigida	rpm
PCNfR	Porcentagem corrigida da velocidade do ventilador	pct
phi	Razão de fluxo de combustível para Ps30	pps/psi

Cada ciclo de voo registrado contém um comprimento de dados e condições de voo diferentes (subida, descida e de cruzeiro). Neste banco de dados as falhas de degradação dos motores são encontrados em dois componentes: HPT e LPT. São dois tipos de falhas que podem ser identificadas nos motores, a saber, uma falha que afeta somente a eficiência da HPT que esta presente nas unidades 2, 5 e 10 e uma mais complexa que prejudica em conjunto a eficiência e o fluxo da LPT e a eficiência da HPT, esses tipos de falhas são identificados nas unidades 16, 18, 20, 11, 14 e 15 (Chao et al., 2021).

Nas Figuras 2, 3, 4 são mostrados as curvas dos nove motores (unidades) em condições de falha de uma degradação anômala da eficiência da HPT (HPT_eff), na eficiência da LPT (LPT_eff) e no fluxo do LPT (LPT_Flow) respectivamente. Cada curva representa a degradação de uma unidade de motor Turbofan (Chao et al., 2021).

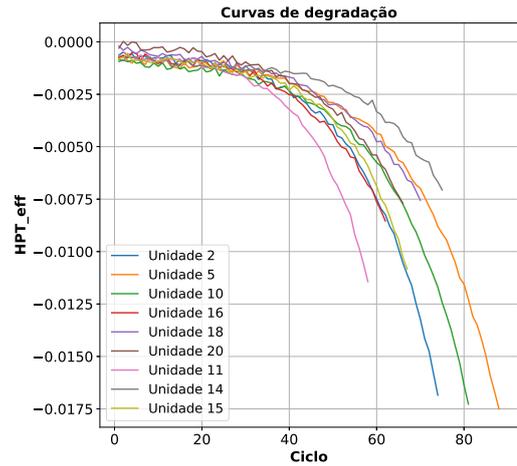


Figura 2. Curvas de degradação com falhas na eficiência da HPT.

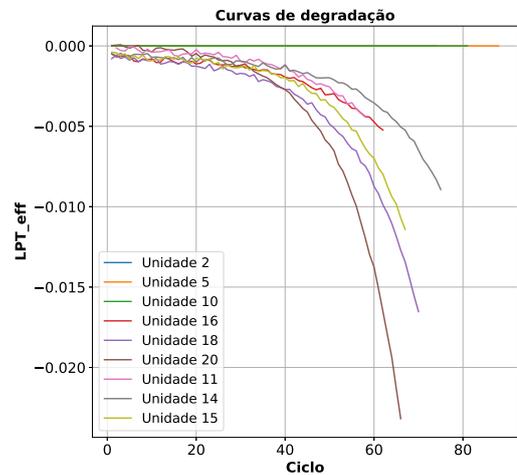


Figura 3. Curvas de degradação com falhas na eficiência da LPT.

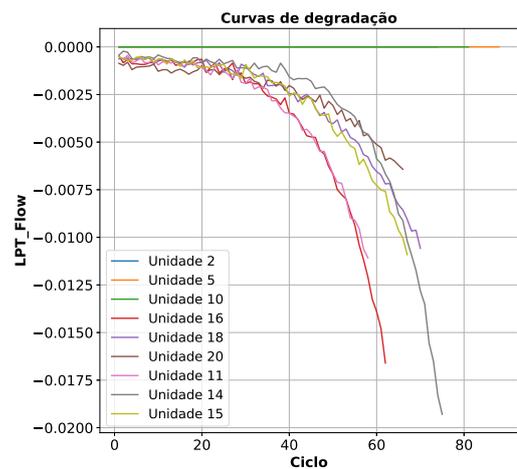


Figura 4. Curvas de degradação com falhas no fluxo da LPT.

Pode-se observar que nas Figuras 3 e 4, como já mencionado anteriormente, não ocorreram falhas na eficiência e no fluxo da LPT nas unidades 2, 5 e 10, por este motivo

as curvas destas unidades não possuem esses tipos de degradação.

Como relatado por Chao et al. (2021) existem 6 unidades utilizadas para treinamento (2, 5, 10, 16, 18 e 20) e 3 unidades para teste (11, 14 e 15). É importante destacar que as unidades 14 e 15 possuem uma operação de voo que difere das demais. Elas operam em altitudes mais baixas e em voos mais curtos. Com isso o conjunto de treinamento contém perfis de voos que não são totalmente representativos para estas duas unidades.

3. TÉCNICAS PARA A MODELAGEM

As técnicas de aprendizado de máquina utilizadas para a modelagem das curvas de degradação das turbinas foram: Árvores de Decisão (AD), Floresta Aleatória (FA) e o *Gradient Boosting* (GB).

Estes algoritmos foram escolhidos pela sua grande aplicabilidade e por suas vantagens: interpretabilidade do modelo, um alto desempenho nos resultados e a capacidade de generalização (Witten and Frank, 2005; Faceli et al., 2011).

3.1 Árvores de Decisão

A AD é um algoritmo aplicado em estimação cujo objetivo é a criação de um modelo capaz de se ajustar o mais próximo possível dos dados.

Para realizar este processo o algoritmo cria regras de decisão no formato de “se” e “então”, cuja representação pode ser feita por um grafo acíclico que lembra a representação de uma árvore (Faceli et al., 2011; Castro and Ferrari, 2016). Na Figura 5 é representado um exemplo de árvore de decisão.

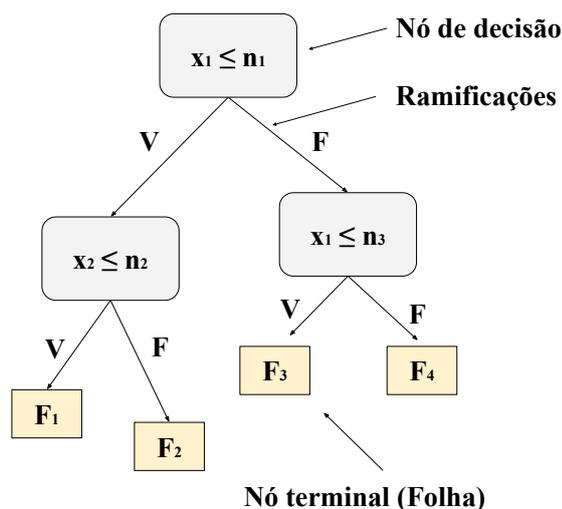


Figura 5. Exemplo de uma árvore de decisão.

A árvore é formada por nós, cada um destes representando uma escolha para a ramificação da árvore. A primeira escolha é chamado nó raiz e dele partirão ramificações da

árvore que podem resultar em outros nós de decisão ou em nós terminais (Folhas) (Faceli et al., 2011).

Este algoritmo tem como diferencial a possibilidade de se visualizar e interpretar as regras de decisão criadas, possibilitando que o analista de dados entenda melhor o problema em que está trabalhando (Castro and Ferrari, 2016). O pseudocódigo da árvore de decisão retirado de Faceli et al. (2011) é apresentado no Algoritmo 1.

Algoritmo 1 Pseudocódigo da AD

Entrada: Conjunto de treinamento $D = \{(x_i, y_i), i = 1, \dots, n\}$
 /* **Função GeraÁrvore(D) */;**

Saída: Árvore de decisão
se Critério de parada (D) = verdadeiro **então**
 Retorna: Rotula o nó folha com a constante que minimiza a função de perda;
fim se
 Escolha o atributo que maximiza o critério de divisão em D;
para Partição de exemplos D_i , baseado nos valores do atributo escolhido **faça**
 Induz uma subárvore $\hat{Árvore}_i = \text{GeraÁrvore}(D_i)$
fim para
Retorna: Árvore contendo um nó de decisão baseado no atributo escolhido, e descendentes $\hat{Árvore}_i$;

Um problema que pode surgir na elaboração deste algoritmo é que dependendo dos dados utilizados para criá-lo, é possível que a árvore se torne complexa e de difícil interpretação.

Outro problema é no sobreajuste que, dependendo do seu processo de elaboração e dos dados, podem ser criadas muitas regras que tornam a árvore superdimensionada (profunda) (Faceli et al., 2011).

O sobreajuste ocorre quando o algoritmo também estima os ruídos (*outliers*) dos dados, desta forma o modelo se especifica tanto para cada dado que acaba se tornando complexo e profundo em excesso prejudicando a generalização. Uma possível solução seria a poda, na qual, é feito a retirada de partes da árvore (sub-nós) que não representam um ganho significativo na sua elaboração (Faceli et al., 2011).

3.2 Floresta Aleatória

Outra opção para tentar diminuir o sobreajuste e conseguir melhores resultados na estimação é a combinação de modelos mais simples de árvores de decisão.

A FA emprega essa opção na criação do seu modelo. Neste algoritmo o banco de dados é dividido em vários subconjuntos de forma aleatória. Com cada subconjunto é criado uma árvore de decisão (Witten and Frank, 2005; Friedman et al., 2009).

Depois que as árvores são criadas ocorre um processo para calcular a média das estimações de cada árvore em um ponto. Ao final desta etapa é criado o modelo da FA (Friedman et al., 2009). Na Figura 6 é representado o processo de criação de uma FA.

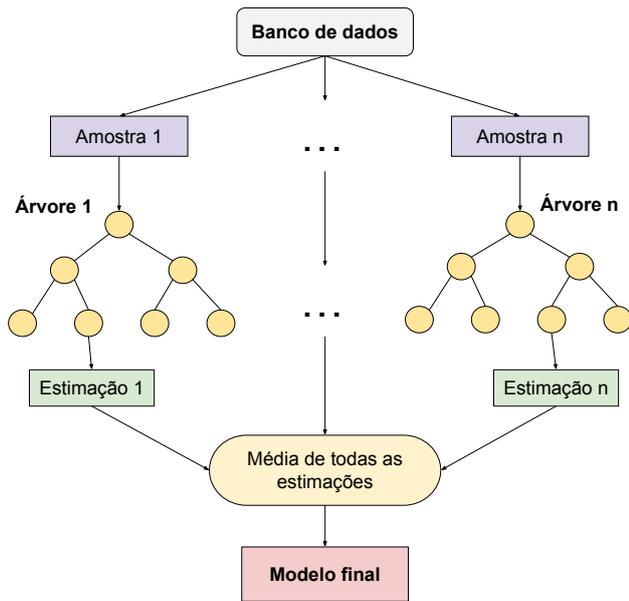


Figura 6. Exemplo da criação de uma FA.

3.3 Gradient Boosting

Outro método de combinação de árvores de decisão é o GB. Neste algoritmo as árvores são criadas de forma sequencial, desta forma, cada uma das árvores tenta se ajustar utilizando os erros dos modelos anteriores. O algoritmo é inicializado com a criação de uma única árvore formada pelo nó terminal, com o avanço das interações, ou seja, a criação de uma nova árvore, o modelo vai se ajustando com o objetivo de melhorar o seu desempenho (Witten and Frank, 2005; Friedman et al., 2009).

Geralmente este tipo de modelo precisa de um número maior de estimadores se comparado com a floresta aleatória, mas em compensação, as árvores criadas no modelo são menos profundas (rasas) que as criadas pelo outro, gerando assim um estimador menos propenso ao sobreajuste (Witten and Frank, 2005; Friedman et al., 2009). Na Figura 7 é mostrado os passos para a elaboração de uma GB.

Na Figura 7, X representa a matriz de atributos e y a variável alvo do problema. O termo y'_1 é o valor estimado utilizado para calcular os erros residuais r_1 . O processo é repetido até que todas as n árvores sejam criadas.

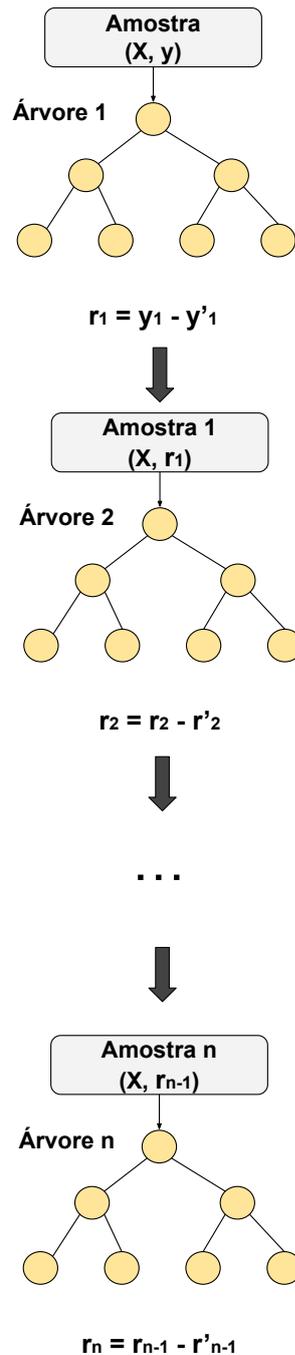


Figura 7. Exemplo da criação de uma GB.

4. METODOLOGIA

Para o desenvolvimento do trabalho foi utilizado o banco de dados do repositório público da NASA disponível através do site <https://www.nasa.gov/content/prognostics-center-of-excellence-data-set-repository>. Na Figura 8 são apresentados os passos empregados para a elaboração deste artigo.

Para a definição e criação dos algoritmos foi empregado a linguagem de programação Python (Van Rossum and Drake Jr, 1995) e a biblioteca *Scikit-Learn* (Buitinck et al., 2013).

Utilizando os dados indicados na base do trabalho de (Chao et al., 2021), cujo o modelo matemático é citado na seção 2 deste artigo, foram definidos quais algoritmos seriam empregados na estimacão dos dados, também foi feito a seleção das colunas que representavam as amostras para se realizar a modelagem. Desta seleção foram retirados os atributos que continuam as amostras dos ciclos e das curvas de degradação impostas na eficiência do HPT, na eficiência do LPT e no fluxo do LPT. Para a modelagem foi utilizado o número de ciclos como entrada e as curvas de degradação como saída para cada caso.

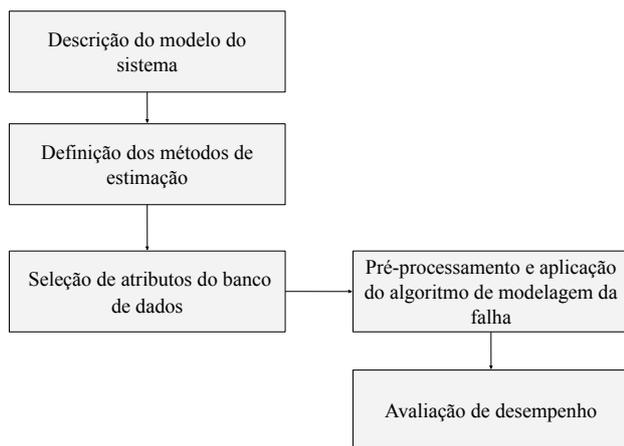


Figura 8. Metodologia utilizado para a criação do trabalho.

No pré-processamento dos dados foi feita uma limpeza das amostras que continham elementos faltantes e ruídos (*outliers*) que poderiam interferir nos resultados dos modelos. Os dados foram divididos em 70% para treinamento e 30% para teste. Posteriormente os algoritmos são treinados e testados.

Por fim, Como forma de avaliação de desempenho, utilizou-se três métricas para definir a eficiência do modelo na estimação dos dados. As métricas são: o Coeficiente de Determinação, o Erro Quadrático Médio e a Raiz do Erro Quadrático Médio.

4.1 Coeficiente de Determinação - R^2

O Coeficiente de Determinação R^2 é usado para estabelecer a adequação de um modelo aos dados. O valor deste coeficiente fica entre 0 e 1, desta forma quanto mais próximo de 1, mais o modelo está ajustado aos dados. Em (2) é mostrado matematicamente este coeficiente:

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

onde y_i representa o valor real, \hat{y}_i é definido como o valor estimado pelo modelo, \bar{y} é a média do valor real e n o número de amostras.

4.2 Erro Quadrático Médio

O Erro Quadrático Médio, do inglês *Mean Squared Error* (MSE), é uma forma de quantificar o quanto um modelo é melhor que o outro. Com isso, quanto mais próximo o erro for de 0, melhor o modelo será (Junior et al., 2017; Castro and Ferrari, 2016). Ele é definido como apresentado em (3):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

4.3 Raiz do Erro Quadrático Médio

A Raiz do Erro Quadrático Médio, do inglês *Root Mean Square Error* (RMSE), também é uma medida de avaliação

de desempenho para casos de estimação e da mesma forma que o MSE, quanto mais próximo seu valor de 0, melhor será o modelo (Castro and Ferrari, 2016).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

5. RESULTADOS

Aplicando a metodologia proposta na seção anterior, foi possível chegar nos resultados expostos nas Tabelas 5, 6 e 7, onde é apresentado a Média (M) e o Desvio Padrão (DP) do R^2 , MSE e RMSE para cada algoritmo e cada tipo de degradação (HPT_eff, LPT_eff e LPT_Flow). Essas medidas (média e desvio padrão) são empregadas para que se tenha uma noção do grau de variação dos resultados encontrados na estimação das curvas para cada algoritmo.

Tabela 5. Média e desvio padrão resultante da estimação das curvas para a degradação na eficiência da HPT.

Modelo	Estatística	R^2	MSE	RMSE
AD	M	0,981	$1,36 \times 10^{-7}$	$3,60 \times 10^{-4}$
	DP	0,005	$6,41 \times 10^{-8}$	$8,65 \times 10^{-5}$
FA	M	0,987	$9,48 \times 10^{-8}$	$2,90 \times 10^{-4}$
	DP	0,007	$7,87 \times 10^{-8}$	$1,03 \times 10^{-4}$
GB	M	0,985	$1,00 \times 10^{-7}$	$3,08 \times 10^{-4}$
	DP	0,005	$4,89 \times 10^{-8}$	$7,19 \times 10^{-5}$

Para as curvas de degradação na eficiência da HPT, as árvores foram criadas de forma que cada nó gera mais dois nós. A profundidade das árvores foi de 4. A floresta aleatória usou 300 estimadores (árvores) em sua criação e foi colocado uma profundidade de 5. O *gradient boosting* foi constituído de 500 estimadores (árvores) e uma profundidade de 3.

Observando os resultados das métricas para a estimação das curvas de degradação na eficiência da HPT (tabela 5), o modelo que realizou a melhor estimação foi a FA, tendo um valor de R^2 de 0,987 e o pior resultado encontrado foi na AD com 0,981. Mesmo assim, os resultados são bem próximos. Já para os valores de MSE e RMSE todos os algoritmos conseguiram resultados bem próximos de 0, mostrando que todos os modelos são aplicáveis para este procedimento. O desvio padrão também obteve uma variação pequena entre os resultados das curvas.

Tabela 6. Média e desvio padrão resultante da estimação das curvas para a degradação na eficiência da LPT.

Modelo	Estatística	R^2	MSE	RMSE
AD	M	0,980	$1,87 \times 10^{-7}$	$3,79 \times 10^{-4}$
	DP	0,007	$1,86 \times 10^{-7}$	$2,09 \times 10^{-4}$
FA	M	0,982	$2,12 \times 10^{-7}$	$3,78 \times 10^{-4}$
	DP	0,007	$2,49 \times 10^{-7}$	$2,62 \times 10^{-4}$
GB	M	0,982	$1,59 \times 10^{-7}$	$3,51 \times 10^{-4}$
	DP	0,007	$1,58 \times 10^{-7}$	$1,91 \times 10^{-4}$

Já para as curvas de degradação com falha na eficiência da LPT, foram utilizados de hiperparâmetros para a AD, uma profundidade de 4 e cada nó gerando mais dois nós. A FA foi feita com 300 árvores e uma profundidade de

5. Para o GB, foi utilizado uma profundidade de 3 e 100 estimadores para sua criação.

No caso dos resultados dos parâmetros para a eficiência da LPT (Tabela 6), observando a métrica R^2 houve um empate entre a FA e o GB com um resultado de 0,982. Já observando pela métrica MSE, o GB conseguiu os melhores resultados por possuir um valor mais próximo de 0. O desvio padrão para o caso do R^2 teve uma variação em todas as métricas de 0,007.

Tabela 7. Média e desvio padrão resultante da estimação das curvas para a degradação no fluxo da LPT.

Modelo	Estatística	R^2	MSE	RMSE
AD	M	0,980	$1,66 \times 10^{-7}$	$3,97 \times 10^{-4}$
	DP	0,008	$7,56 \times 10^{-8}$	$9,51 \times 10^{-5}$
FA	M	0,980	$1,82 \times 10^{-7}$	$4,05 \times 10^{-4}$
	DP	0,007	$1,07 \times 10^{-7}$	$1,31 \times 10^{-4}$
GB	M	0,981	$1,65 \times 10^{-7}$	$3,92 \times 10^{-4}$
	DP	0,007	$8,22 \times 10^{-8}$	$1,05 \times 10^{-4}$

Por fim, para as curvas de degradação com a falha no fluxo da LPT tiveram os seguintes hiperparâmetros: uma profundidade de 4 e cada nó sendo dividido em dois outros nós para a AD, uma profundidade de 4 e possuindo 550 estimadores a floresta aleatória e a GB com uma profundidade de 2 e com 150 estimadores em sua constituição.

Na estimação das curvas para a degradação no fluxo da LPT (Tabela 7) o algoritmo com melhor resultado foi a GB, já o pior, foi um empate entre a AD e a FA. O desvio padrão também, como no caso anterior, para a métrica R^2 teve uma variação de 0,007 para a FA e o GB e de 0,008 para a AD.

Nas Figuras 9, 10 e 11 são apresentados os dados reais e os dados estimados pelos algoritmos que obtiveram os melhores resultados na classificação para cada tipo de degradação.

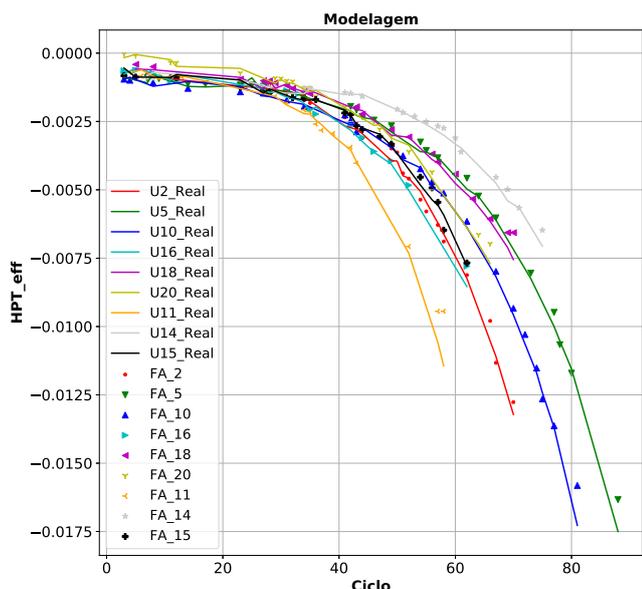


Figura 9. Dados reais e estimados para a FA nas curvas de degradação na eficiência da HPT.

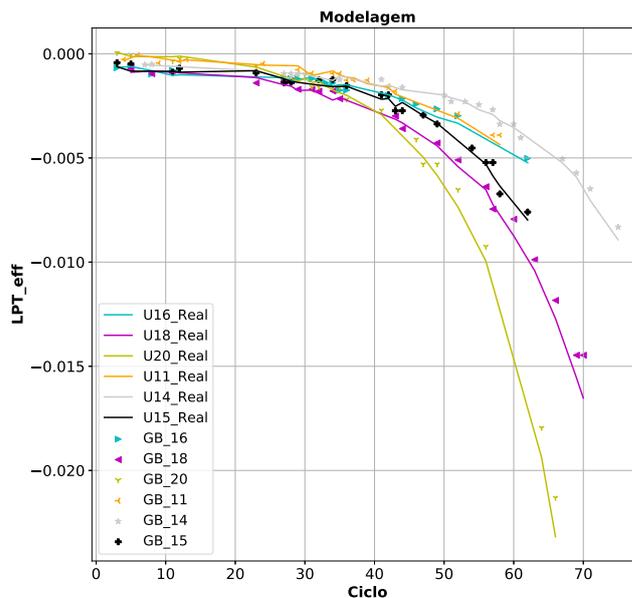


Figura 10. Dados reais e estimados para o GB nas curvas de degradação da eficiência da LPT.

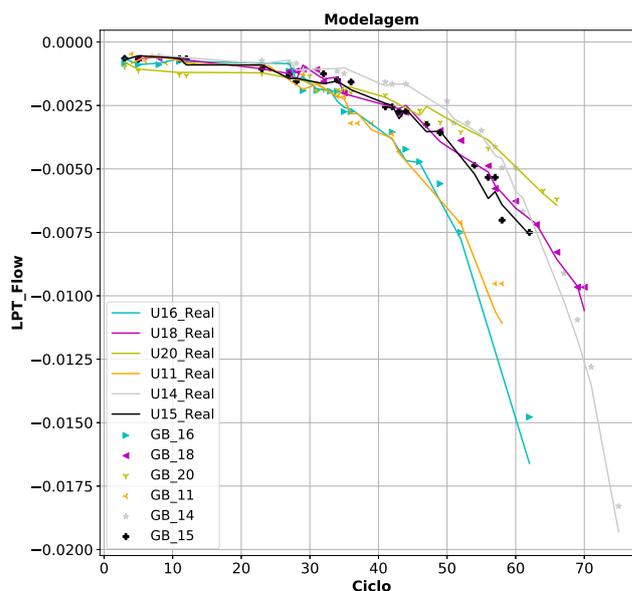


Figura 11. Dados reais e estimados para o GB nas curvas de degradação do fluxo da LPT.

Nas Figuras 9, 10 e 11 os dados reais estão representados pelo número da unidade e cada uma possuindo cores diferentes, já os dados previstos possuem a mesma cor dos dados reais, mas estão apresentados através de símbolos. Observando a Figura 9 pode-se perceber que a FA, para as curvas de degradação na eficiência da HPT, conseguiu uma estimação próxima do real possuindo apenas alguns pontos mais distantes dos dados.

Já nas Figuras 10 e 11 a estimação dos dados também foram bem próximas, possuindo poucos pontos que se destoam dos moldes das curvas. Neste dois casos não houve falhas na eficiência e no fluxo do LPT nas unidades 2, 5 e 10, por este motivo essas unidades não estão representadas na figura.

6. CONCLUSÃO

O trabalho realizou um comparativo de desempenho entre três algoritmos que utilizam as árvores de decisão em sua composição para realizar o modelo de degradação em turbinas de aeronaves em condições reais de voo. Os resultados obtiveram valores de R^2 superiores a 0,98 e valor de MSE e RMSE pequenos e próximos de 0.

Para as curvas de degradação na eficiência da HPT o melhor resultado encontrado foi utilizando a FA. Já para as curvas de degradação da eficiência e no fluxo da LPT os melhores resultados foram encontrados utilizando o GB.

Mesmo com valores abaixo dos outros algoritmos, as árvores criadas não ficaram muito profundas, possibilitando o estudo e interpretação das regras de decisão. Desta forma, como explicitado pelos resultados, o uso destas técnicas trazem duas grandes vantagens: a interpretabilidade das regras criadas pela árvore de decisão e os ótimos desempenhos alcançados pelos modelos.

Em trabalhos futuros espera-se utilizar novas técnicas de aprendizado de máquina e realizar um comparativo com modelos estatísticos de estimação. Também deseja-se realizar todo o processo de prognóstico da vida útil destas turbinas, com isso o sistema terá capacidade de fornecer previsões da degradação e conseqüentemente ajudar no estabelecimento de políticas de manutenção preventiva destes equipamentos.

AGRADECIMENTOS

Este trabalho foi realizado com o apoio da Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP) através da concessão de bolsa de estudo de Mestrado e auxílio através do projeto BP4-0172-00277.01.00/20 BPI 02/2020

REFERÊNCIAS

- Ameneh, F.S., Prakash, Y.O., and Haitao, L. (2017). A review on degradation modelling and its engineering applications. *International Journal of Performability Engineering*.
- Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., and Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*.
- Castro, L.N.D. and Ferrari, D.G. (2016). *Introdução à mineração de dados: conceitos básicos, algoritmos e aplicações*, volume 1. Saraiva, São Paulo.
- Chao, M.A., Kulkarni, C., Goebel, K., and Fink, O. (2021). Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics. *Data*.
- Faceli, K., Lorena, A.C., Gama, J., and de Leon Ferreira de Carvalho, A.C.P. (2011). *Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina*, volume 1. LTC, Rio de Janeiro.
- Frederick, D., DeCastro, J., and Litt, J. (2007). User's guide for the commercial modular aero-propulsion system simulation (c-mapss). *NASA Technical Manuscript*.

- Friedman, J.H., Tibshirani, R., and Hastie, T. (2009). *The Elements of Statistical Learning*, volume 1. Springer.
- Junior, W.R.L., Almeida, V.M., and Martins, S.A.M. (2017). Identificação de um motor/gerador cc por meio de modelos polinomiais autorregressivos e redes neurais artificiais.
- Kang, R., Gong, W., and Chen, Y. (2020). Model-driven degradation modeling approaches: Investigation and review. *Chinese Journal of Aeronautics*.
- Li, Z., Goebel, K., and Wu, D. (2019). Degradation modeling and remaining useful life prediction of aircraft engines using ensemble learning. *Journal of Engineering for Gas Turbines and Power-transactions of The Asme*.
- Lopes, C.B. (2012). *Determinação da política ótima de manutenção em sistemas reparáveis sujeitos a manutenções imperfeitas* 2012. 78 f. Mestrado em engenharia de produção, Universidade Federal de Minas Gerais, Belo Horizonte.
- Saxena, A., Goebel, K., Simon, D., and Eklund, N. (2008). Damage propagation modeling for aircraft engine run-to-failure simulation. *International Conference on Prognostics and Health Management*.
- Sogut, M.Z., Yalcin, E., and Karakoc, T.H. (2017). Assessment of degradation effects for an aircraft engine considering exergy analysis. *Energy. Advanced Energy Technologies in Aviation*.
- Van Rossum, G. and Drake Jr, F.L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.
- Witten, I.H. and Frank, E. (2005). *Data mining : practical machine learning tools and techniques*, volume 2. Elsevier.
- Zhang, B., Zheng, K., Huang, Q., Feng, S., Zhou, S., and Zhang, Y. (2020). Aircraft engine prognostics based on informative sensor selection and adaptive degradation modeling with functional principal component analysis. *Sensors*.