

Deep Learning aplicado em previsão de falhas em turbinas de aeronaves e equipamentos industriais

Gabriel Vinicius Boin Freitas¹

¹Centro de Ciências Tecnológicas – Universidade Estadual de Santa Catarina (UDESC)
Rua Paulo Malschitzki, 200 Zona Industrial Norte, CEP: 89.219-710 – Joinville – SC – Brazil

`gvb.freitas@edu.udesc.br`

Abstract. *This study investigates the application of deep learning techniques for predicting the Remaining Useful Life (RUL) of aircraft turbines and large industrial equipment. Using a hybrid approach that combines Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and Transformer layers, the proposed model aims to capture both local features and long-term temporal dependencies in sensor data. The data used is sourced from NASA's C-MAPSS dataset. The custom loss function, custom score, is designed to penalize prediction errors more severely when the prediction is later than the actual RUL. The model was evaluated on four data subsets, showing promising performance in RUL prediction, with evaluation metrics including R^2 , MAPE, MAE, MSE, RMSE, and the custom score. Results suggest that the combination of CNN, LSTM, and Transformer is effective for predictive maintenance, with potential for future improvements through hyperparameter tuning, integration of additional data, and validation in real-world scenarios.*

Key words: *RUL: Remaining Useful Life - Deep Learning - Turbine - Anomaly detection - Predictive maintenance - CNN - LSTM - Transformer*

Resumo. *Este estudo investiga a aplicação de técnicas de aprendizado profundo para a previsão da Vida Útil Restante (RUL) em turbinas de aeronaves e grandes equipamentos industriais. Utilizando uma abordagem híbrida que combina Redes Neurais Convolucionais (CNN), Redes Neurais Recorrentes (LSTM) e camadas Transformer, o modelo proposto visa capturar tanto as características locais quanto as dependências temporais de longo prazo nos dados de sensores. Os dados utilizados são provenientes do Conjunto de Dados de Prognósticos de Motores a Jato (C-MAPSS) fornecido pela NASA. A função de perda personalizada, custom score, foi projetada para penalizar mais severamente os erros de previsão onde a previsão é posterior ao RUL real. O modelo foi avaliado em quatro subconjuntos de dados, mostrando desempenho promissor na previsão de RUL, com métricas de avaliação incluindo R^2 , MAPE, MAE, MSE, RMSE e a pontuação customizada. Resultados sugerem que a combinação de CNN, LSTM e Transformer é eficaz para a manutenção preditiva, com potencial para melhorias futuras através de ajustes de hiperparâmetros, integração de dados adicionais e validação em cenários reais.*

1. Introdução

Nos últimos anos, o uso de técnicas de aprendizado profundo (deep learning) tem crescido significativamente em várias áreas, incluindo a indústria aeronáutica e a indústria de petróleo. Uma aplicação promissora dessa tecnologia é a previsão de falhas em turbinas de aeronaves e em grandes equipamentos industriais, como turbinas de termelétricas, grandes bombas e outros equipamentos utilizados na indústria do petróleo. [Fentaye et al. 2021, Dai et al. 2024]

O problema da previsão de falhas em turbinas de aeronaves e em grandes equipamentos industriais é de grande importância devido aos altos custos associados à manutenção corretiva e ao tempo de inatividade dos equipamentos. Além disso, a detecção precoce de falhas pode ajudar a evitar acidentes graves, reduzir os custos de manutenção e melhorar a eficiência operacional. [Fentaye et al. 2021]

Uma turbina aeronáutica é um tipo de motor a jato que converte a energia cinética do ar em movimento em energia mecânica rotacional. O processo começa com a entrada de ar na turbina, onde é comprimido em alta pressão por uma série de estágios de compressão. Em seguida, o ar comprimido é misturado com combustível e queimado em uma câmara de combustão, gerando uma rápida expansão dos gases quentes. Esses gases de alta temperatura são então expelidos por uma série de turbinas rotativas, que estão ligadas a um eixo central. A energia cinética resultante do fluxo de gases em alta velocidade através das turbinas é convertida em energia mecânica, que pode ser usada para movimentar hélices ou gerar eletricidade em um avião.

O funcionamento de turbinas utilizadas em termelétricas é essencialmente o mesmo que o de turbinas aeronáuticas. A diferença está no uso final da energia mecânica gerada. Em uma termelétrica, a energia mecânica rotacional do eixo da turbina é usada para girar um gerador elétrico, convertendo-a em eletricidade que pode ser distribuída para a rede elétrica.

As termelétricas são de interesse da indústria do petróleo porque muitas vezes são alimentadas por combustíveis fósseis, como petróleo e gás natural. Além disso, a eletricidade gerada por termelétricas é frequentemente usada em processos de extração, refino e transporte de petróleo, bem como em instalações de produção offshore, onde a energia elétrica é necessária para operar equipamentos de perfuração, bombeamento e tratamento de água.

Os principais modos de falha em turbinas aeronáuticas são erosão, fadiga do material, fissuras, desgaste dos rolamentos, corrosão, desbalanceamento, vazamento de óleo, problemas de vedação, degradação do revestimento térmico. [Dai et al. 2024]

O objetivo geral deste trabalho é desenvolver e avaliar modelos de aprendizado profundo para prever falhas em turbinas de aeronaves e em grandes equipamentos industriais com alta precisão e antecedência. Para alcançar esse objetivo, objetivos específicos serão abordados, sendo coletar e pré-processar dados de sensores de turbinas de aeronaves, utilizando a base de dados da NASA disponível no "Prognostics Center of Excellence Data Repository", utilizar modelos de autoencoder para identificar padrões de assinatura de bom funcionamento nos dados e detectar anomalias que possam indicar falhas iminentes nos equipamentos, aplicar modelos de previsão de séries temporais, como particle filter, para prever o remaining useful life (RUL) dos equipamentos, treinar e otimizar

os modelos desenvolvidos usando conjuntos de dados históricos utilizando a base de dados da NASA disponível no "Prognostics Center of Excellence Data Repository" e avaliar o desempenho dos modelos em termos de precisão, recall e F1-score, bem como compará-los com métodos tradicionais de previsão de falhas.

A metodologia de pesquisa adotada neste estudo incluirá uma revisão abrangente da literatura sobre previsão de falhas em turbinas de aeronaves e em grandes equipamentos industriais, aquisição e pré-processamento de dados utilizando a base da NASA, desenvolvimento e treinamento de modelos de aprendizado profundo, e avaliação do desempenho dos modelos. Espera-se que os resultados deste estudo contribuam significativamente para o avanço da segurança e eficiência operacional na indústria aeronáutica e na indústria de petróleo.

2. Revisão de conceitos e trabalhos relacionados

A previsão da vida útil remanescente (RUL) é uma área crítica para a manutenção preditiva e redução de custos operacionais. Diversos estudos recentes têm explorado técnicas avançadas de aprendizado profundo para melhorar a precisão dessas previsões. Ragab et al. (2020) introduziram o modelo ATS2S, que utiliza técnicas de sequência para sequência com um mecanismo de atenção, otimizando simultaneamente a perda de reconstrução e a perda de previsão de RUL [Ragab et al. 2020]. Este modelo foi destacado por sua eficácia em superar 13 métodos estado-da-arte em quatro conjuntos de dados reais, particularmente em cenários complexos com múltiplas condições operacionais. Soni et al. (2020) exploraram o uso de redes neurais convolutivas (CNN) para prever a RUL, comparando com abordagens baseadas em Long Short-Term Memory (LSTM). Os resultados indicaram que o modelo CNN proposto alcançou uma precisão de 91,74% a 94,16% em testes experimentais, superando as abordagens baseadas em LSTM [Soni et al. 2020].

Outro estudo notável é o de Wang et al. (2020a), que utilizou pares de redes adversárias para facilitar o treinamento de redes neurais, com a estabilidade assegurada por um algoritmo de amostragem baseado em RMSprop. Este método demonstrou uma melhoria significativa na previsão de RUL em comparação com métodos existentes de fusão de dados [Wang et al. 2020a]. Em termos de revisões e análises, Namuduri et al. (2020) revisaram os algoritmos de aprendizado profundo usados para manutenção preditiva, destacando a importância dos sensores eletroquímicos e ópticos para monitoramento industrial [Namuduri et al. 2020]. Wang et al. (2020b) discutiram como técnicas de aprendizado profundo, como Autoencoders, DBN, CNN e RNN, oferecem soluções poderosas para a previsão de RUL, superando os desafios associados aos modelos baseados em física [Wang et al. 2020b]. Pang et al. (2021) realizaram uma análise detalhada das técnicas de deep learning para detecção de anomalias, destacando seu potencial para melhorar a precisão na identificação de anomalias em dados complexos [Pang et al. 2021].

Além desses modelos, Al-Khazraji et al. (2022) desenvolveram um modelo híbrido de Autoencoder e Redes de Crença Profunda (DBN) para a previsão de RUL de motores de aeronaves. Este modelo se destacou por sua capacidade de extrair características abstratas dos dados de entrada e prever a série temporal da RUL, demonstrando superioridade em comparação com outras abordagens de aprendizado profundo [Al-Khazraji et al. 2022]. Li et al. (2022) propuseram um modelo que combina CNN, LSTM e o módulo de atenção de bloco convolucional (CBAM) para prever a RUL de

motores de aeronaves, alcançando uma redução de 17,4% no RMSE e uma melhoria de 25,9% no score em comparação a outros métodos estado-da-arte [Li et al. 2022]. Hu et al. (2023) apresentaram modelos de fusão baseados em Transformer que integram redes LSTM e CNN para melhorar a precisão da previsão de RUL de motores aeroespaciais, resultando em uma melhoria de até 66,53% em RMSE e 84,86% nas métricas de pontuação [Hu et al. 2023].

Finalmente, Khalaf e Ali (2023) exploraram a influência de estratégias de manutenção na performance e eficiência das turbinas a gás, enfatizando a importância da manutenção baseada no estado para melhorar a confiabilidade e reduzir os custos do ciclo de vida [Khalaf and Ali 2023].

3. Desenvolvimento da proposta de solução

A previsão da Vida Útil Restante (RUL) é uma tarefa essencial para a manutenção preditiva de motores de aeronaves. Este estudo utiliza uma abordagem híbrida que combina redes convolucionais (CNN), redes neurais recorrentes (LSTM) e camadas Transformer para prever o RUL com base em dados multivariados de sensores. A combinação dessas técnicas visa capturar tanto as características locais quanto as dependências temporais de longo prazo nos dados.

Os dados utilizados são provenientes do Conjunto de Dados de Prognósticos de Motores a Jato (C-MAPSS) fornecido pela NASA. Este conjunto de dados contém múltiplas séries temporais multivariadas divididas em quatro subconjuntos: FD001, FD002, FD003 e FD004, cada um com diferentes condições operacionais e modos de falha. Cada série temporal representa o ciclo operacional de um motor, registrando 26 variáveis em cada ciclo, incluindo configurações operacionais e medições de sensores. No subconjunto FD001, há 100 trajetórias de treino e 100 de teste, sob uma única condição operacional e um modo de falha. O subconjunto FD002 possui 260 trajetórias de treino e 259 de teste, sob seis condições operacionais e um modo de falha. O FD003 inclui 100 trajetórias de treino e 100 de teste, sob uma única condição operacional e dois modos de falha. Finalmente, o FD004 contém 248 trajetórias de treino e 249 de teste, sob seis condições operacionais e dois modos de falha.

Cada motor começa sua operação em condições normais e desenvolve uma falha durante a série temporal. No conjunto de treinamento, a falha progride até a falha total do sistema, enquanto no conjunto de teste, a série temporal termina algum tempo antes da falha total. O objetivo é prever o número de ciclos operacionais restantes antes da falha total no conjunto de teste.

Os dados são carregados e pré-processados utilizando a biblioteca *pandas*. Durante o carregamento, colunas completamente vazias são removidas. Os dados são então normalizados usando *StandardScaler* para escalar as medições dos sensores. Os dados de treinamento são preparados em sequências de comprimento fixo, definidas como 40 ciclos neste estudo, para capturar a dinâmica temporal dos sensores. Cada sequência é emparelhada com o RUL correspondente, calculado como o número de ciclos restantes até a falha. Os dados de teste são preparados de forma semelhante, com preenchimento (*padding*) para sequências menores que o comprimento fixo.

O modelo híbrido proposto é composto por uma combinação de camadas CNN,

LSTM e Transformer. Primeiramente, as camadas convolucionais (CNN) utilizam 64 filtros com tamanho de kernel 3 na primeira camada e 128 filtros na segunda camada, com pooling para reduzir a dimensionalidade. As camadas CNN são responsáveis por extrair características locais dos dados temporais. Em seguida, duas camadas LSTM com 200 unidades cada são usadas para capturar dependências temporais de longo prazo. A primeira camada retorna sequências, enquanto a segunda retorna a última saída. Finalmente, uma camada Transformer é utilizada para melhorar a capacidade do modelo de focar em diferentes partes da sequência de entrada. O bloco Transformer possui 200 dimensões de embedding, 2 cabeças de atenção e 200 dimensões na rede feed-forward. O modelo é compilado utilizando o otimizador `adam` e uma função de perda customizada baseada na fórmula fornecida para a pontuação customizada.

A função de perda personalizada, `custom_score`, é projetada para penalizar de forma diferenciada os erros de previsão. A função penaliza mais severamente os erros onde a previsão é posterior ao RUL real, em comparação com os erros onde a previsão é anterior ao RUL real. A fórmula da função de perda é dada por:

$$d = \hat{y} - y$$

onde \hat{y} é o valor previsto e y é o valor real do RUL. A pontuação é calculada da seguinte forma:

$$\text{score}(d) = \begin{cases} \exp\left(-\frac{d}{13}\right) - 1, & \text{se } d < 0 \\ \exp\left(\frac{d}{10}\right) - 1, & \text{se } d \geq 0 \end{cases}$$

Assim, a função de perda final é a soma das pontuações para todos os pares de previsões e valores reais:

$$\text{custom_score}(y, \hat{y}) = \sum_{i=1}^n \text{score}(d_i)$$

onde n é o número total de exemplos.

O modelo é treinado utilizando `early_stopping` para prevenir overfitting. O `early_stopping` monitora a perda na validação e restaura os melhores pesos após 10 épocas sem melhoria. O treinamento é realizado por 250 épocas com um tamanho de lote de 40, e uma divisão de validação de 20% dos dados de treinamento. As métricas de avaliação incluem R^2 , MAPE, MAE, MSE, RMSE e a pontuação customizada, fornecendo uma visão abrangente do desempenho do modelo na previsão de RUL.

Os resultados do treinamento são visualizados através de gráficos do histórico de perda durante as épocas de treinamento. As previsões de RUL são comparadas com os valores reais utilizando scatter plots, permitindo uma análise detalhada das previsões do modelo. Além disso, um histograma de erro das previsões de RUL é plotado para visualizar a distribuição dos erros. A combinação de CNN, LSTM e Transformer permite capturar efetivamente características locais e globais dos dados de sensores, resultando em previsões precisas de RUL. Esta abordagem híbrida demonstrou ser eficaz na tarefa

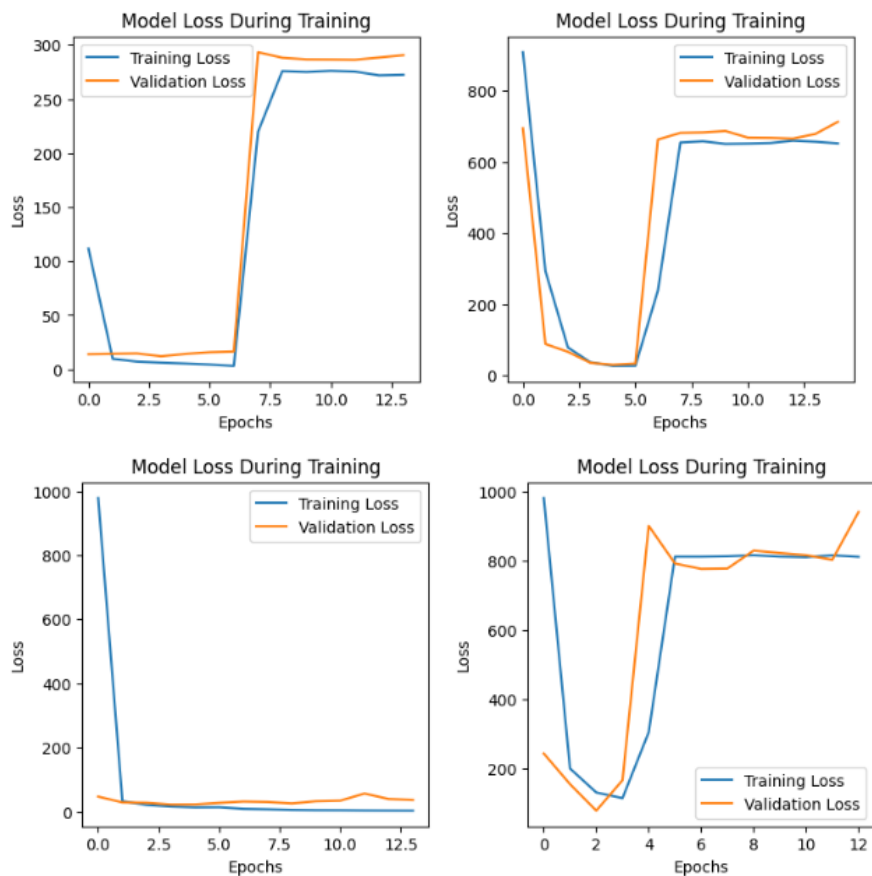


Figure 1. Curvas de perda durante o treinamento e validação para os subconjuntos FD001, FD002, FD003 e FD004.

de manutenção preditiva, fornecendo uma ferramenta valiosa para a gestão de motores de aeronaves.

Link para código: <https://github.com/gabrielfea/CMAPSSData>

4. Análise de Resultados e Conclusão

A análise dos resultados foi conduzida em quatro subconjuntos de dados do Conjunto de Dados de Prognósticos de Motores a Jato (C-MAPSS). As figuras a seguir apresentam os resultados do treinamento, os histogramas de erro de previsão, as comparações entre previsões e valores reais de RUL, e as métricas de avaliação para os quatro subconjuntos: FD001, FD002, FD003 e FD004.

A Figura 1 mostra as curvas de perda durante o treinamento e a validação para os quatro subconjuntos de dados. Observa-se que o modelo conseguiu convergir rapidamente nas primeiras épocas para a maioria dos conjuntos de dados, mas há variações significativas na estabilidade das curvas de perda.

A Figura 2 apresenta os histogramas dos erros de previsão para os quatro subconjuntos de dados. A distribuição dos erros fornece uma visão clara da precisão do modelo. No geral, os erros estão concentrados em torno de zero, mas há diferenças na dispersão dos erros entre os diferentes conjuntos de dados.

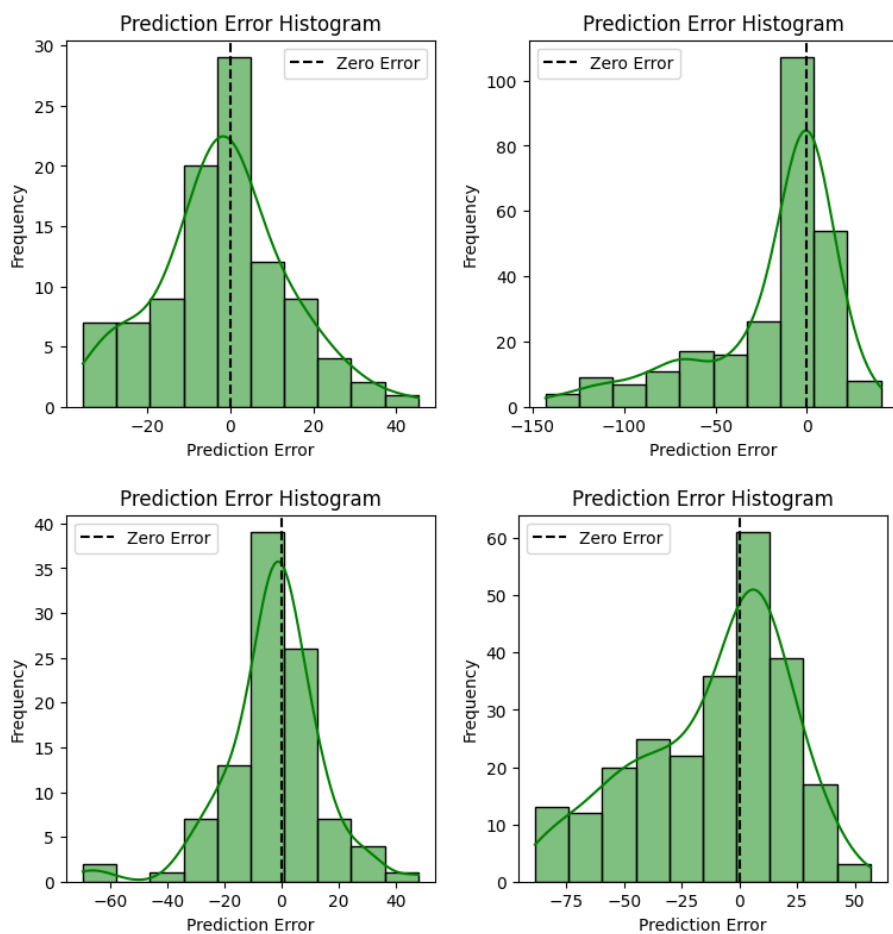


Figure 2. Histogramas dos erros de previsão para os subconjuntos FD001, FD002, FD003 e FD004.

A Figura 3 mostra a comparação entre as previsões de RUL e os valores reais para os quatro subconjuntos. As previsões de RUL estão geralmente próximas aos valores reais, embora existam discrepâncias notáveis em alguns pontos, indicando áreas para melhorias no modelo.

A Tabela 4 resume as métricas de avaliação para os quatro subconjuntos de dados. As métricas incluem R^2 , MAPE, MAE, MSE, RMSE e a pontuação customizada. As métricas indicam um bom desempenho do modelo em termos de precisão das previsões de RUL, com variações entre os diferentes conjuntos de dados.

Os resultados apresentados indicam que a abordagem híbrida utilizando CNN, LSTM e Transformer é eficaz na previsão da Vida Útil Restante (RUL) de motores a jato. A função de perda personalizada, `custom_score`, desempenha um papel crucial ao penalizar mais severamente os erros onde a previsão é posterior ao RUL real. Esta penalização diferenciada é essencial para aplicações de manutenção preditiva, onde falhas não detectadas a tempo podem ter consequências significativas.

Apesar do bom desempenho geral, as discrepâncias observadas em algumas previsões e a variabilidade nas curvas de perda sugerem que há espaço para melhorias. Ajustes finos nos hiperparâmetros e a exploração de técnicas adicionais de regularização podem ajudar a melhorar ainda mais a estabilidade e a precisão do modelo.

5. Trabalhos Futuros

Embora os resultados obtidos com a abordagem híbrida de CNN, LSTM e Transformer tenham mostrado um bom desempenho na previsão da Vida Útil Restante (RUL) de motores a jato, há várias direções promissoras para trabalhos futuros que podem melhorar ainda mais a eficácia do modelo e sua aplicação em cenários reais.

Uma área potencial de melhoria é o ajuste fino dos hiperparâmetros do modelo. A exploração de técnicas como a busca em grade (grid search) ou a otimização bayesiana pode ajudar a identificar configurações de hiperparâmetros que proporcionem melhor desempenho. Além disso, a implementação de mecanismos avançados de regularização, como Dropout variacional e normalização por lotes (batch normalization), pode aumentar a estabilidade do treinamento e reduzir o risco de overfitting.

Outra direção importante é a incorporação de dados adicionais, como dados de manutenção histórica, registros de falhas anteriores e condições ambientais. A integração desses dados pode fornecer informações contextuais valiosas que podem melhorar a precisão das previsões de RUL. O uso de técnicas de fusão de dados (data fusion) pode ser explorado para combinar essas diferentes fontes de informação de maneira eficaz.

A explicabilidade e a interpretabilidade dos modelos de previsão de RUL são críticas para a sua aceitação em aplicações industriais. Trabalhos futuros podem focar no desenvolvimento de métodos para explicar as decisões do modelo, como mapas de atenção (attention maps) para camadas Transformer, e técnicas de interpretação local como LIME (Local Interpretable Model-agnostic Explanations).

Finalmente, a validação do modelo em cenários reais de operação de motores a jato é essencial para confirmar sua utilidade prática. Estudos de caso e experimentos de campo podem ser conduzidos para avaliar o desempenho do modelo em condições

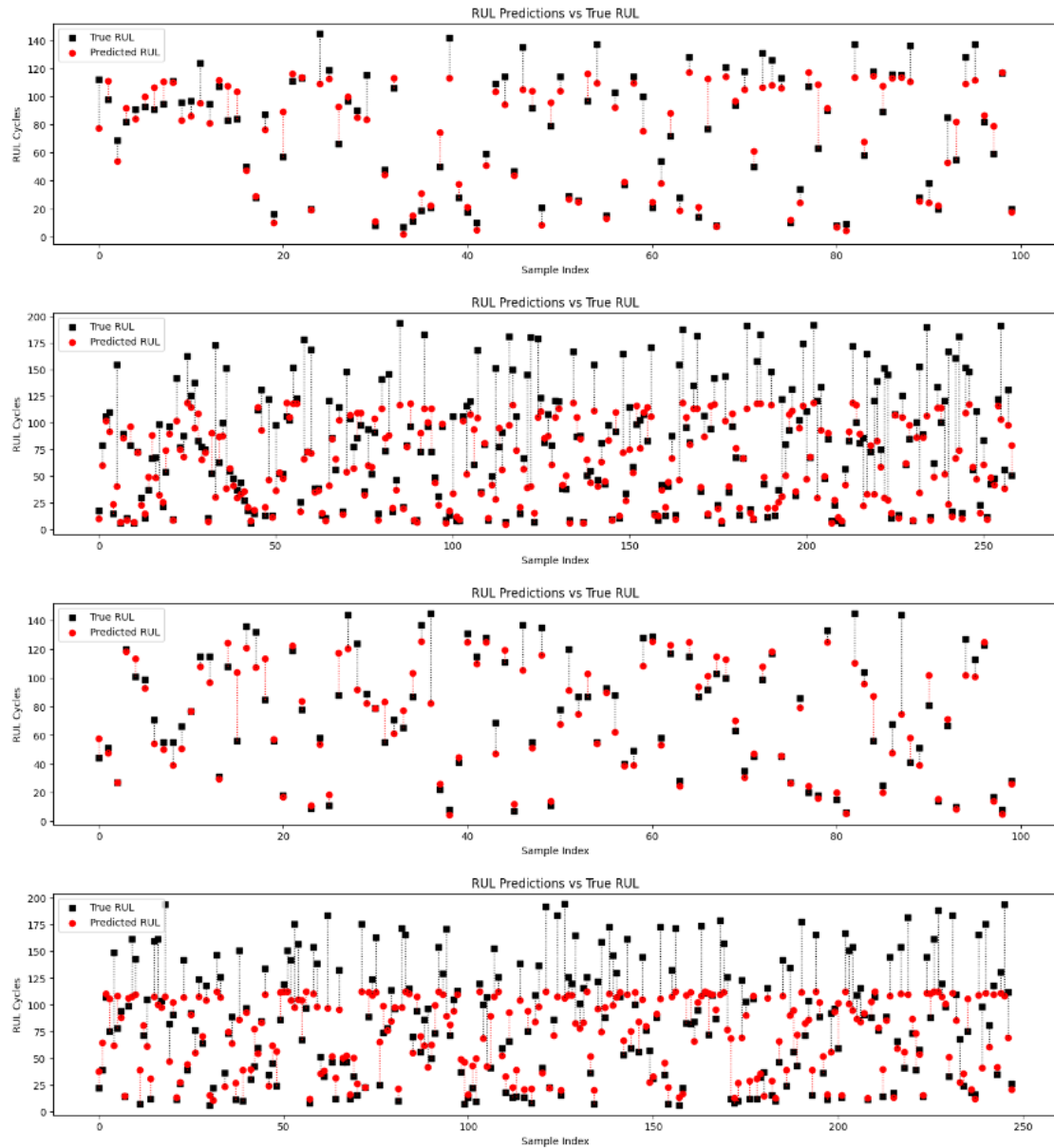


Figure 3. Comparação entre previsões de RUL e valores reais para os subconjuntos FD001, FD002, FD003 e FD004.

	Metric	Value		Metric	Value
0	R ²	0.8536382922345618	0	R ²	0.3948005386483817
1	MAPE	0.19610056098937556	1	MAPE	0.27629648255796196
2	MAE	12.097572898864746	2	MAE	25.66715567636674
3	MSE	252.74758374421907	3	MSE	1750.332889596474
4	RMSE	15.898037103498629	4	RMSE	41.83697992920227
5	Score	414.9918212890625	5	Score	274158.59375

	Metric	Value		Metric	Value
0	R ²	0.8338831745568672	0	R ²	0.605266665339102
1	MAPE	0.16969951336356232	1	MAPE	0.46319523985439476
2	MAE	11.457517681121827	2	MAE	25.997780530683457
3	MSE	284.65739339897124	3	MSE	1173.4806359694921
4	RMSE	16.87179283297929	4	RMSE	34.25610363087857
5	Score	688.6266479492188	5	Score	12701.3955078125

Figure 4. Métricas de avaliação para os subconjuntos FD001, FD002, FD003 e FD004.

operacionais reais e ajustar o modelo conforme necessário para atender às necessidades específicas da indústria.

A implementação prática do modelo em sistemas de manutenção preditiva pode ser explorada, incluindo a integração com sistemas de monitoramento de saúde de motores (Engine Health Monitoring Systems - EHMS). O desenvolvimento de uma interface de usuário amigável e a automação do processo de coleta e análise de dados são passos importantes para a adoção industrial da tecnologia.

Em resumo, há várias oportunidades empolgantes para expandir e aprimorar o trabalho realizado até agora, avançando na direção de sistemas de manutenção preditiva mais precisos, interpretáveis e aplicáveis em cenários reais.

References

- Al-Khazraji, H., Nasser, A. R., Hasan, A. M., Al Mhdawi, A. K., Al-Raweshidy, H., and Humaidi, A. J. (2022). Aircraft engines remaining useful life prediction based on a hybrid model of autoencoder and deep belief network. *IEEE Access*, 10:82156–82163.
- Dai, S., Zhang, X., and Luo, M. (2024). A novel data-driven approach for predicting the performance degradation of a gas turbine. *Energies*, 17(4):781.
- Fentaye, A. D., Zaccaria, V., and Kyprianidis, K. (2021). Aircraft engine performance monitoring and diagnostics based on deep convolutional neural networks. *Machines*, 9(12):337.
- Hu, Q., Zhao, Y., and Ren, L. (2023). Novel transformer-based fusion models for aero-engine remaining useful life estimation. *IEEE Access*.
- Khalaf, K. D. and Ali, O. M. (2023). Effect of the maintenance strategy on the performance and efficiency of the gas turbine unit: A review. *Journal Européen des Systèmes Automatisés*, 56(2):259.

- Li, H., Wang, Z., and Li, Z. (2022). An enhanced cnn-lstm remaining useful life prediction model for aircraft engine with attention mechanism. *PeerJ Computer Science*, 8:e1084.
- Namuduri, S., Narayanan, B. N., Davuluru, V. S. P., Burton, L., and Bhansali, S. (2020). Deep learning methods for sensor based predictive maintenance and future perspectives for electrochemical sensors. *Journal of The Electrochemical Society*, 167(3):037552.
- Pang, G., Shen, C., Cao, L., and Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. *ACM computing surveys (CSUR)*, 54(2):1–38.
- Ragab, M., Chen, Z., Wu, M., Kwoh, C.-K., Yan, R., and Li, X. (2020). Attention sequence to sequence model for machine remaining useful life prediction. *arXiv preprint arXiv:2007.09868*.
- Soni, H., Kansara, A., and Joshi, T. (2020). Predictive maintenance of gas turbine using prognosis approach. *International Research Journal of Engineering and Technology*, 7(6):4683–4691.
- Wang, F., Du, J., Zhao, Y., Tang, T., and Shi, J. (2020a). A deep learning based data fusion method for degradation modeling and prognostics. *IEEE Transactions on Reliability*, 70(2):775–789.
- Wang, Y., Zhao, Y., and Addepalli, S. (2020b). Remaining useful life prediction using deep learning approaches: A review. *Procedia manufacturing*, 49:81–88.