,

UNIVERSIDADE FEDERAL DO PARÁ
CAMPUS UNIVERSITÁRIO DE TUCURUÍ
FACULDADE DE ENGENHARIA ELÉTRICA

WILLIAM MACHADO DA SILVA

PROGNÓSTICO DE FALHAS BASEADO EM REDES NEURAIS CONVOLUCIONAIS: Uma análise para estimar a vida útil remanescente de turbinas turbofan

Trabalho de Conclusão de Curso

Tucuruí - PA
2021

UNIVERSIDADE FEDERAL DO PARÁ
CAMPUS UNIVERSITÁRIO DE TUCURUÍ
FACULDADE DE ENGENHARIA ELÉTRICA

WILLIAM MACHADO DA SILVA

PROGNÓSTICO DE FALHAS BASEADO EM REDES NEURAIS CONVOLUCIONAIS: Uma análise para estimar a vida útil remanescente de turbinas turbofan

Trabalho de Conclusão de Curso apresentado à Faculdade de Engenharia Elétricada Universidade Federal do Pará, Campus Universitário de Tucuruí, para obtenção do grau de Engenheiro Eletricista.
Orientador: Prof. PhD. Rafael Suzuki Bayma

Tucuruí - PA
2021

WILLIAM MACHADO DA SILVA

PROGNÓSTICO DE FALHAS BASEADO EM REDES NEURAIS CONVOLUCIONAIS: Uma análise para estimar a vida útil remanescente de turbinas turbofan

Trabalho de Conclusão de Curso apresentado à Faculdade de Engenharia Elétricada Universidade Federal do Pará, Campus Universitário de Tucuruí, para obtenção do grau de Engenheiro Eletricista.
Aprovada em: 03/06/2021

Prof. PhD. Rafael Suzuki Bayma (Orientador) - UFPA

Prof. Dr. Cleison Daniel Silva (Examinador) - UFPA

Prof. Dr. Raphael Barros Teixeira (Examinador) - UFPA

Tucuruí - PA
2021

AGRADECIMENTOS

Em primeiro lugar, a Deus, que fez com que meus objetivos fossem alcançados, durante todos os meus anos de estudos.

A minha mãe Elma Ferreira, meu pai Antônio Carlos e meu irmão Wallace Machado, que me incentivaram nos momentos difíceis e compreenderam a minha ausência enquanto eu me dedicava à realização deste trabalho.

Aos amigos, Elenildo Cássio, Gabrielle Helene e Marcelo Monteiro que sempre estiveram ao meu lado, pela amizade incondicional e pelo apoio demonstrado ao longo de todo o período de tempo em que me dediquei a este trabalho.

Ao professor Rafael Suzuki Bayma, por ter sido meu orientador e ter desempenhado tal função com dedicação e amizade.

Às pessoas com quem convivi ao longo desses anos de curso, que me incentivaram e que certamente tiveram impacto na minha formação acadêmica e pessoal.

A todos os amigos que de alguma forma fizeram parte dessa jornada eu agradeço com um forte abraço.

RESUMO

Este trabalho propõe um método baseado em aprendizado profundo para estimar a vida útil remanescente (RUL do inglês, *Remaining Useful Life*) de turbinas turbofan. É apresentada uma arquitetura de rede neural convolucional (CNN do inglês, *Convolutional Neural Network*). O método proposto é aplicado a um conjunto de dados da NASA (do inglês, *National Aeronautics and Space Administration*) para a estimativa RUL, variando o conjunto de sensores e a forma de degradação das turbinas em dois modelos: RUL linear e RUL linear por partes. Os resultados experimentais são comparados entre si e com outros métodos encontrados na literatura, as comparações mostram que o método proposto RUL linear por partes apresenta desempenho comparável aos outros trabalhos publicados em termos da raiz do erro quadrático médio (RMSE do inglês, *Root Mean Square Error*) e uma métrica específica de competições de prognósticos.

Palavras-chave— Aprendizado profundo, vida útil remanescente, rede neural convolucional, RUL linear e RUL linear por partes.

ABSTRACT

This work proposes a deep learning method for estimating the remaining useful life (RUL) of turbofan turbines. It presents a convolutional neural network (CNN) architecture. The proposed method is applied to a NASA data set for the RUL estimate, by varying the set of sensors and the form of turbine degradation in two models: linear RUL and piecewise linear RUL. Experimental results are compared with each other and with other methods found in the literature; the results show that the proposed piecewise linear RUL method exhibits similar performance found in other published works in terms of root mean squared error and a specific error metric used in competitions.

Keywords— Deep learning, remaining useful life, convolutional neural network, linear RUL and piecewise linear RUL.

Dedico este trabalho a Deus por nortear minha vida,
Aos meus pais e irmão pelo exemplo, incentivo, amor e carinho,
Aos meus amigos pela convivência, apoio e atenção nos momentos alegres e tristes.

LISTA DE FIGURAS

Figura 1 – Ilustração da vida útil remanescente (RUL), na qual com a sua estimação permite a prevenção de falhas, de modo que uma manutenção efetiva possa ser realizada em tempo de corrigir falhas que estão próximas de acontecer. O sinal de degradação de uma ou mais variáveis pode servir como base para estimar a RUL de um determinado sistema. 17
Figura 2 – Ilustração hierárquica das abordagens de prognóstico, nas abordagens próximas ao topo da pirâmide observa-se um aumento do custo relativo de desenvolvimento e também um aumento do desempenho desejado, em contrapartida há perda de generalidade da abordagem, o que a torna mais específica em sua aplicação (SIMEÓN, 2015). 20
Figura 3 – Ilustração da estrutura hierárquica AI vs ML vs DL. 24
Figura 4 – Ilustração de um neurônio biológico (à esquerda) com a direção do fluxo do sinal e uma sinapse (à direita). 25
Figura 5 – Esquema de representação de um neurônio artificial, onde a soma ou corpo da célula é representada por uma composição de dois módulos, o primeiro é uma junção aditiva, somatório dos estímulos (sinais de entrada) multiplicado pelo seu fator excitatório (pesos sinápticos), e posteriormente uma função de ativação, que definirá com base nas entradas e pesos sinápticos, qual será a saída do neurônio. O axônio é aqui representado pela saída (y_k) obtida pela aplicação da função de ativação (ACADEMY, 2021). 26
Figura 6 – Ilustração de um MLP com três camadas. 27
Figura 7 – Ilustração do processo de treinamento de uma ANN. 27
Figura 8 – Ilustração do processo de treinamento, onde a seta azul aponta na direção da retropropagação e a seta vermelha aponta na direção da propagação direta. . 29
Figura 9 – A configuração do ancestral das CNNs, o “LeNet”. 30
Figura 10 – A configuração da primeira CNN profunda, a “AlexNet”. 31
Figura 11 – Ilustração de uma convolução 1D para séries temporais multivariadas, os kernel são caracterizados pelos retângulos vermelhos e amarelos. 33
Figura 12 – Diagrama simplificado da turbina simulada no C-MAPSS. 36
Figura 13 – Layout mostrando vários módulos e suas conexões conforme modelado na simulação 36
Figura 14 – Ilustração das variáveis selecionadas do conjunto 05 na turbina 1. 40
Figura 15 – Medições do sensor 2 da turbina 1 normalizado. 41
Figura 16 – Janela deslizante para um único sensor. 42
Figura 17 – Funções alvo RUL. 43
Figura 18 – Proposta de arquitetura CNN 1D para estimativa da RUL. 44
Figura 19 – Distribuição dos erros das estimativas da RUL para CNN 1D com RUL linear 49

Figura 20 – Distribuição dos erros das estimativas do RUL para CNN 1D com RUL linear por partes 51

LISTA DE TABELAS

Tabela 1 – Detalhes dos conjuntos de dados (Simulado no C-MAPSS) 37
Tabela 2 – Exemplo da estrutura dos dados FD001. 38
Tabela 3 – Saídas do C-MAPSS para medir a resposta do sistema. 38
Tabela 4 – Resultados da CNN 1D com RUL linear 48
Tabela 5 – Resultados da CNN 1D com RUL linear por partes 50
Tabela 6 – Resultados de desempenhos de outros autores. 52

LISTA DE ABREVIATURAS E SIGLAS

AI	*Artificial Intelligence*
ANN	*Artificial Neural Networks*
API	*Application Programming Interface*
ATOMIC	*Automotive PredicTOr Maintenance In Cloud*
C-MAPSS	*Commercial Modular Aero-Propulsion System Simulation*
CNN	*Convolutional Neural Network*
DL	*Deep Learning*
EOL	*End-of-Life*
ETTF	*Estimated Time To Failure*
GUI	*Graphical User Interface*
HPC	*High Pressure Compressor*
HPT	*High Pressure Turbine*
IVHM	*Integrated Vehicle Health Management*
LPC	*Low Pressure Compressor*
LPT	*Low Pressure Turbine*
ML	*Machine Learning*
MLP	*Multi Layer Perceptron*
MSE	*Mean Square Error*
MSTAR	*Moving and Stationary Target Acquisition and Recognition*
NASA	*National Aeronautics and Space Administration*
PCA	*Principal Component Analysis*
PHM	*Prognostic Health Management*
PROMISA	*Prognostic Module for Intelligent System Analysis*
ReLu	*Rectified Linear Unit*
RMSE	*Root Mean Squared Error*

RTF *run-to failure*

RUL *Remaining Useful Life*

SUMÁRIO

1 INTRODUÇÃO 14
1.1 Considerações gerais 14
1.2 Objetivos 15
1.2.1 Objetivo geral 15
1.2.2 Objetivos específicos 16
1.3 Estrutura do documento 16
2 FUNDAMENTAÇÃO TEÓRICA 17
2.1 Vida útil remanescente 17
2.2 Prognóstico de falhas 18
2.3 Abordagens de prognóstico 19
2.3.1 Abordagens baseadas em experiências 19
2.3.2 Abordagens baseadas em modelos físicos 21
2.3.3 Abordagens baseadas em dados 21
2.4 Aquisição e processamento de dados 22
2.4.1 Aquisição de dados 22
2.4.2 Processamento de dados 23
2.4.2.1 Extração de características 23
2.4.2.2 Seleção de características 23
2.5 Aprendizagem profunda 24
2.6 Redes neurais artificiais 25
2.6.1 Perceptron 26
2.6.2 Treinamento supervisionado 27
2.6.2.1 Função de ativação 28
2.6.3 Rede Neural Convolucional 29
2.6.3.1 Diferenças CNN 1D x 2D 31
2.6.3.2 Camada de convolução 31
2.6.3.3 Camada de pooling 33
2.6.3.4 Camada totalmente conectada 34
3 METODOLOGIA 35
3.1 Descrição dos dados 35
3.1.1 C-MAPSS 35
3.2 Estrutura dos dados 35
3.3 Seleção do subconjunto de dados 37
3.4 Pré-processamento dos dados 39
3.4.1 Seleção das variáveis 39
3.4.2 Normalização 41
3.4.3 Janela deslizante 41

3.4.4 Função alvo RUL 42
3.4.4.1 RUL linear 42
3.4.4.2 RUL linear por partes 43
3.5 Arquitetura da CNN 43
3.6 Métricas de desempenho 44
3.6.1 Pontuação S 44
3.6.2 Raiz do erro quadrático médio 45
4 RESULTADOS E DISCUSSÕES 46
4.1 CNN 1D 46
4.2 Avaliação de desempenho 46
4.2.1 CNN 1D com RUL linear 47
4.2.2 CNN 1D com RUL linear por partes 48
4.2.3 Resultado dos testes 50
4.3 *Benchmarking* 52
4.4 Síntese do capítulo 52
5 CONCLUSÕES E TRABALHOS FUTUROS 53
5.1 Conclusões 53
5.2 Trabalhos futuros 54
REFERÊNCIAS 55

1 INTRODUÇÃO

1.1 Considerações gerais

Ao longo das últimas décadas, a consciência empresarial tem se baseado em seu capital intangível, no conhecimento e na expertise de seus colaboradores. Se uma organização não produz um produto ou serviço de forma eficiente, um concorrente o fará. A necessidade de equilibrar custos, garantir a qualidade máxima e colocar o produto no mercado o mais rápido possível significa que as organizações devem ser flexíveis o suficiente para lidar com o mercado em evolução. Uma área que afeta todos esses aspectos é a gestão e manutenção de ativos (LEVITT, 2003).

Como os orçamentos operacionais costumam ser os primeiros a serem cortados, não faz mais sentido empregar a mentalidade de "esperar para quebrar". Os executivos dizem que o maior impacto nas operações é a falha de ativos críticos (LEVITT, 2003).

Sendo assim, as tecnologias de prognóstico são cruciais na manutenção baseada em condições para diversas áreas de aplicação, como manufatura, aeroespacial, automotiva, industrial e etc. Ao obter a degradação do funcionamento de um equipamento, as tecnologias de prognóstico estimam o desempenho futuro de um subsistema ou um componente para fazer a estimativa da RUL. Se for possível prever com precisão quando um motor irá falhar, então é plausível tomar decisões de manutenção informadas com antecedência para evitar desastres, reduzindo assim o custo de manutenção, bem como agilizar as atividades operacionais (BABU; ZHAO; LI, 2016; WANG et al., 2008).

Porém, os prognósticos baseados em dados enfrentam o desafio da falta de conjuntos de dados de execução até a falha (RTF do inglês, *Run To Fail*). A aquisição de dados reais de progressão de falha do sistema são normalmente caros ou a maioria não são devidamente instrumentalizados para a coleta de dados e aqueles que conseguem coletar dados para uma frota de sistemas tendem a manter os dados privados (SAXENA et al., 2008). Na maioria das vezes, os conjuntos de dados não apresentam a evolução temporal da falha, e existem poucos repositórios públicos que disponibilizam dados de RTF, por exemplo o repositório de prognósticos da Administração Nacional da Aeronáutica e Espaço (NASA do inglês, *National Aeronautics and Space Administration*) é uma coleção de conjuntos de dados que foram doados por várias universidades, agências ou empresas que possuem pesquisa em prognóstico, ou seja, que podem ser usados para o desenvolvimento de algoritmos e estudo de caso. Na maioria das vezes, são séries temporais que variam de algum estado nominal a um estado de falha (NASA, 2021).

Este trabalho descreve o desenvolvimento de um algoritmo baseado em dados para prever a RUL de um sistema complexo à medida que ele se deteriora do estado inicial desconhecido para a falha. Diversos conjuntos de dados de série temporal multivariados foram fornecidos

como parte do problema da degradação de uma frota de turbinas turbofan, simulada utilizando o software de simulação de sistema de aeropropulsão modular comercial (C-MAPSS do inglês, *Commercial Modular Aero-Propulsion System Simulation*). Cada conjunto de dados é representado por uma única unidade do sistema que possui desgaste inicial arbitrário e variação de fabricação desconhecida. Os conjuntos de dados incluem três configurações operacionais e vinte e uma medições de sensores contaminados com ruído. Conjuntos de "treinamento"e "teste"são fornecidos. O conjunto de treinamento contém exemplos de unidades que funcionam até a falha, enquanto os conjuntos de teste terminam algum tempo antes da falha. Os conjuntos de teste são usados para avaliar a precisão das soluções apresentadas (NASA, 2021).

Este conjunto de dados foi utilizado na primeira conferência internacional sobre prognósticos e gestão da daúde (PHM do inglês, *Prognostics And Health Management*), na qual tem como objetivo:

- Promover o desenvolvimento, crescimento e reconhecimento de prognósticos e gestão de saúde (PHM) como uma disciplina de engenharia;
- Apoiar a educação em PHM desenvolvendo currículos de ensino padrão na área;
- Participar e facilitar a colaboração científica internacional em conexão com o avanço ou aplicação do PHM;
- Estabelecer, apresentar, administrar ou organizar conferências, workshops, seminários e cursos de PHM;
- Publicar publicações revisadas por pares concernentes ou relacionadas ao PHM e suas aplicações;
- Estabelecer, desenvolver ou adotar padrões, métodos e métricas em PHM.

A solução apresentada aqui foi desenvolvida em linguagem de programação Python. E apresenta uma arquitetura de CNN, sendo este um algoritmo que funciona de maneira similar aos neurônios do cérebro humano, tornando-se capaz de modelar a estimativa de vida útil remanescente de um sistema.

1.2 Objetivos

1.2.1 Objetivo geral

Este trabalho tem como principal objetivo, propor uma abordagem para prognóstico de vida útil remanescente de turbinas turbofan baseada em dados históricos usando redes neurais convolucionais.

1.2.2 Objetivos específicos

Pretende-se alcançar os seguintes objetivos específicos:

- Apresentar uma arquitetura de estimativa da RUL baseado em dados históricos;
- Implementar o estudo de caso utilizando dados de degradação de turbinas turbofan adquiridos do repositório de prognóstico da NASA;
- Comparar duas funções alvo RUL, avaliando qual apresenta melhor desempenho;
- Utilizar cinco conjuntos de sensores baseados em trabalhos publicados, para avaliar os resultados;
- Comparar os resultados obtidos com alguns algoritmos de prognósticos publicados, entre 2013 a 2020.

1.3 Estrutura do documento

Este documento foi estruturado em cinco capítulos.

Capítulo 1: Este capítulo é a introdução ao trabalho onde se apresenta o objetivo geral, os objetivos específicos a serem alcançados, assim como a estrutura do documento.

Capítulo 2: Apresenta a fundamentação teórica com uma revisão da literatura sobre prognóstico de falhas, abordagens, os algoritmos mais comuns e descreve os fundamentos das redes neurais convolucionais.

Capítulo 3: Apresenta o estudo de caso mais a metodologia adotada para estimativa da RUL.

Capítulo 4: Mostra os resultados e as discussões para a abordagem desenvolvida.

Capítulo 5: Apresentam-se as conclusões dos resultados e as sugestões de trabalhos futuros sobre o tema apresentado.

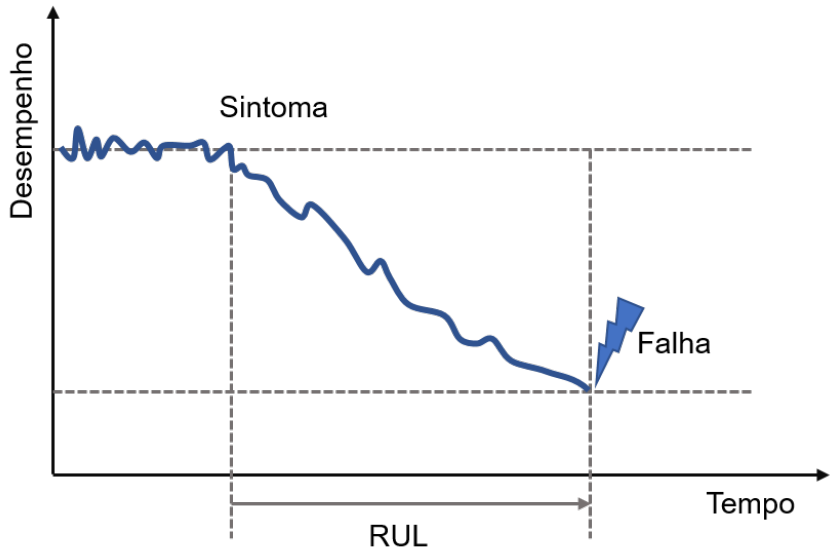
2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo apresentam-se os conceitos necessários para a compreensão do processo de prognóstico da vida útil remanescente (RUL do inglês, *Remaining Useful Life*). As principais aplicações, vantagens e desvantagens de cada um dos prognósticos de falhas são abordados.

2.1 Vida útil remanescente

A RUL é definida como o tempo remanescente, em termos de horas de operação, ciclos e etc., que o equipamento ou sistema continuará operando dentro de suas especificações de projeto, também é conhecida como tempo estimado até a falha (ETTF do inglês, *Estimated Time To Failure*) (SIMEÓN, 2015). A RUL refere-se também ao tempo restante (desde a aparição de um sintoma) até que um determinado limiar de falha seja atingido, como mostrado na Figura 1. Sikorska, Hodkiewicz e Ma (2011) definem o prognóstico de falha como a estimação da RUL de um equipamento.

Figura 1 – Ilustração da vida útil remanescente (RUL), na qual com a sua estimação permite a prevenção de falhas, de modo que uma manutenção efetiva possa ser realizada em tempo de corrigir falhas que estão próximas de acontecer. O sinal de degradação de uma ou mais variáveis pode servir como base para estimar a RUL de um determinado sistema.



Fonte: Elaborada pelo autor.

Conforme indicado por Bengtsson (2004) o desenvolvimento atual em tecnologias para manutenção baseada na condição está caminhando para o prognóstico (isto é, um diagnóstico de situação futura). Thurston e Lebold (2001) definem como a principal função do algoritmo de prognóstico projetar o estado de saúde atual do equipamento no futuro, levando em consideração estimativas de perfis de uso futuros. A camada de prognósticos pode relatar estado de saúde em um momento futuro, ou pode estimar o RUL de um ativo, dado seu perfil de uso projetado.

As técnicas de prognóstico fornecem a capacidade de prever o RUL de um componente ou subsistema com falha. A consistência na estimação de RUL é fundamental para a redução dos custos de manutenção. As medidas de prognóstico podem auxiliar nas decisões de manutenção, na determinação do tempo de início de uma manutenção, e na estimação do tempo para a substituição de peças ou equipamentos degradados, evitando dessa forma manutenções desnecessárias e falhas inesperadas. Nos últimos anos houve um interesse significativo em propostas de metodologias para prognóstico de RUL de equipamentos ou sistemas utilizando diferentes abordagens e algoritmos (SIMEÓN, 2015).

2.2 Prognóstico de falhas

De acordo com Söderholm (2005), falha é definido como: "Término da capacidade de um item para executar uma função exigida". Uma falta, por outro lado, é definida como: "Estado de um item caracterizado pela incapacidade de realizar uma função exigida, excluindo a incapacidade durante a prevenção, manutenção ou outras ações planejadas, ou por falta de recursos externos". Portanto, uma falha é um evento, enquanto uma falta é um estado. Falhas em plantas industriais podem gerar muitos inconvenientes. Todinov (2006) lista diversos possíveis problemas resultantes de falha do sistema, todos os quais podem incorrer em custos elevados: tempo e volume de produção perdida, despejo de produtos químicos nocivos para o meio ambiente, perda de clientes, pagamentos de garantia, custo de mobilização de emergência, custo de seguro e multas relativas a legislação.

Os métodos de previsão têm desempenhado um papel cada vez mais importante na análise de falhas combinando modelagem por computador, métodos de análise de modo/mecanismo. A durabilidade agora pode ser por exemplo, prevista para algum desgaste comum, mecanismos de falha ou fadiga da junta do equipamento. No entanto as condições de teste são frequentemente constituídas por perfis cíclicos predeterminados, que em geral são diferentes dos perfis reais que possuem condições de carga, nível de carga e frequência irregulares ao longo do tempo. Obviamente, isso faz com que os métodos de predição que são baseados em condições constante predeterminadas não sejam adequadas para estimar a vida útil restante dos componentes sob as condições reais de aplicação, e é aqui que o prognóstico pode desempenhar um papel importante (BAILEY et al., 2010).

Os equipamentos ou máquinas e seus componentes em geral passam por um processo de degradação antes de falhar completamente (LEE et al., 2006). A estimação de RUL é realizada monitorando a degradação do equipamento. O processo de degradação pode ser acompanhado por um sistema de prognóstico, com o propósito de monitorar a evolução de um ou mais tipos de falha, e estimar a RUL, uma vez que uma condição de falha seja detectada, isolada e identificada. A estimação da RUL é baseada em três abordagens: em modelos, em dados e em experiências, além de modelos híbridos que surgem a partir da combinação das abordagens

anteriores (SIMEÓN, 2015).

2.3 Abordagens de prognóstico

Na literatura técnica recente, uma variedade de aplicações de prognóstico tem sido relatadas, sendo a maioria exemplo de aplicações projetadas especificamente conforme o equipamento, tais como: motores de indução (BARBOSA et al., 2020), transformadores de potência (ARAÚJO et al., 2007), plantas industriais (TEIXEIRA et al., 2014) e etc. De acordo com Simeón (2015) as pesquisas sobre prognósticos de falhas têm empregado um vasto número de técnicas em diversas áreas do conhecimento, tais como análise de regressão, previsão de séries temporais, inteligência artificial, sistemas especialistas, lógica fuzzy, etc .

Segundo Lee et al. (2006) apesar do progresso alcançado, muitas questões fundamentais ainda precisam ser abordadas:

- A maioria das abordagens de prognóstico desenvolvidas são de aplicação ou equipamento específico;
- Atualmente, os métodos geralmente se concentram em resolver o problema de previsão de falha. Ferramentas para avaliar o desempenho do sistema avaliação e previsão de degradação não foram bem desenvolvidos;
- Muitos algoritmos de previsão desenvolvidos foram demonstrados apenas em um ambiente de laboratório, sem validações em plantas industriais.

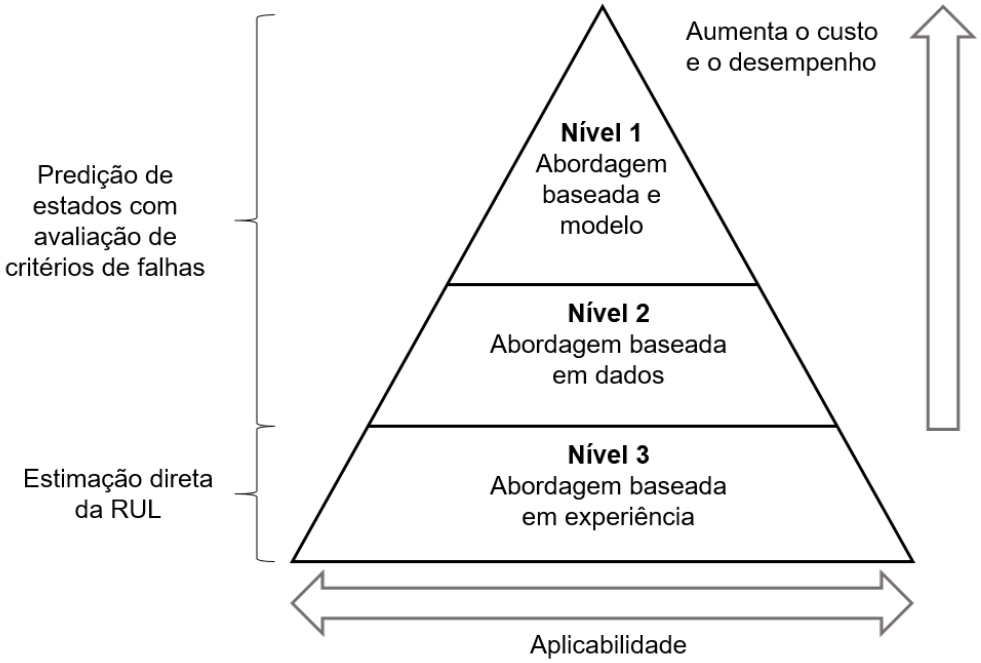
As diferentes técnicas podem ser classificadas segundo Vachtsevanos et al. (2006) em três categorias, as quais utilizam três níveis de prognóstico: abordagens baseadas em modelos, baseadas em dados, e baseadas em experiências, como mostrado na Figura 2.

2.3.1 Abordagens baseadas em experiências

Segundo Liao e Köttig (2014) modelos baseados na experiência correlacionam conhecimento especializado e experiência de engenharia (frequentemente documentada como regras SE-ENTÃO diretamente de especialistas do domínio) com a situação observada para inferir RUL a partir de medidas ou eventos históricos. Sistemas especialistas podem ser considerados modelos baseados na experiência porque é necessário o conhecimento do domínio.

As regras definidas em sistemas especialistas são em geral inteligíveis e de fácil leitura. No entanto, os sistemas especialistas dependem de regras especializadas exatas e completas, e pode sofrer do problema de explosão combinacional quando o número de entradas e saídas aumentam (LIAO; KÖTTIG, 2014).

Figura 2 – Ilustração hierárquica das abordagens de prognóstico, nas abordagens próximas ao topo da pirâmide observa-se um aumento do custo relativo de desenvolvimento e também um aumento do desempenho desejado, em contrapartida há perda de generalidade da abordagem, o que a torna mais específica em sua aplicação (SIMEÓN, 2015).



Fonte: Adaptado de Simeón (2015)

Para avaliar a saúde e estimar a RUL dos equipamentos de grande porte, pode ser necessário implementar centenas ou milhares de regras de produção, o uso de modelos baseados na experiência é representativo de como um especialista humano resolve um problema. No entanto, modelos baseados na experiência dependem muito da capacidade do especialista de domínio para especificar as regras representando todas as características do sistema.

Biagetti e Sciubba (2004) apresentaram a implantação de um sistema especialista capaz de monitorar, por meio de um procedimento de controle inteligente de saúde, o desempenho instantâneo de uma planta de cogeração. O sistema especialista denominado de módulo prognóstico para análise inteligente de sistema (PROMISA do inglês, *Prognostic Module for Intelligent System Analysis*). Gerando, em tempo real e de forma diretamente útil ao gerente da planta, informações sobre a existência e gravidade das falhas, previsões sobre o histórico de tempo futuro das falhas detectadas e prováveis e sugestões de como controlar o problema. A base de conhecimento (fatos e regras) é agrupada - os blocos de conhecimento pertencem a componentes individuais. Uma interface gráfica de usuário (GUI do inglês, *Graphical User Interface*) permite ao usuário interrogar PROMISA sobre suas regras, procedimentos, classes e objetos, e sobre seu caminho de inferência. O artigo também apresenta os resultados de alguns testes de simulação.

2.3.2 Abordagens baseadas em modelos físicos

Segundo Luo et al. (2003) a abordagem baseada em modelos é derivada diretamente de primeiros princípios e da compreensão dos mecanismos físicos. Além disso, as mudanças na saída do modelo, tendem a ter um significado físico direto (ou facilmente traduzível).

No entanto, os modelos baseados na física podem não estar disponíveis para sistemas complexos devido à falta de compreensão de todos os modos de falhas e comportamentos sob uma variedade de condições operacionais. Além disso, um modelo baseado em física é muitas vezes construído de elemento por elemento, que constitui o equipamento. Portanto, geralmente não é aplicável a um sistema diferente sem uma quantidade significativa de esforço (SIKORSKA; HODKIEWICZ; MA, 2011).

A principal vantagem das abordagens baseadas em modelos é a sua capacidade para incorporar descrições físicas fundamentais do equipamento monitorado. Portanto, a precisão das abordagens baseadas em modelos pode ser significativamente maior do que a das abordagens baseadas em dados e baseadas em experiências (SIMEÓN, 2015). Conforme ilustrado na Figura 2, espera-se que as técnicas deste modelo forneçam a previsão de vida útil mais precisa entre todas as técnicas. Este método é sugerido para ser usado onde alta confiabilidade é essencial e atualmente bem-sucedida (BAILEY et al., 2010).

Djeziri et al. (2020), propõem a utilização do conhecimento físico dos componentes para construir modelos, vinculando esses recursos às variáveis mensuráveis, em seguida, os modelos físicos são utilizados para gerar índices de saúde cuja evolução pode ser estimada e prevista online, e os dados de confiabilidade usados para inicializar os modelos de tendência dos índices de saúde. Os resultados experimentais, obtidos para estimar o RUL do transistor de efeito de campo, mostram a eficácia do método proposto.

2.3.3 Abordagens baseadas em dados

Segundo Bailey et al. (2010) técnicas baseadas em dados usam as informações disponíveis e históricas para derivar estatísticas e estimar a RUL do equipamento. Os algoritmos são baseados em modelos aprendidos diretamente a partir dos dados, em vez de um modelo feito com base em perícia humana. Para usar técnicas orientadas a dados, as variáveis precursoras precisam ser identificadas primeiro e monitoradas continuamente durante o uso. Se houver uma significativa mudança nas observações atuais em comparação com a tendência observada com o produto saudável, isso significa que falhas podem estar ocorrendo.

A estimação da RUL é realizada por um algoritmo de prognóstico, depois de detectado um desvio nas condições de operação obtém-se como resultado a progressão das características do sistema até a interseção entre os dados extrapolados e um limiar de falha estabelecido a priori.

De acordo com Liao e Köttig (2014) os modelos baseados em dados são derivados apenas de dados históricos, e são aplicáveis quando os dados são suficientemente abundantes. A saída

gerada a partir do monitoramento tendem a fornecer informações mais concisas do que as obtidas por experiência ou feedback de manutenção. No entanto, a maioria dos resultados dos modelos orientados por dados não são fáceis de explicar ou de receber um significado físico.

Petrillo et al. (2020) desenvolveram um sistema de gestão integrada da saúde do veículo (IVHM do inglês, *Integrated Vehicle Health Management*) para lidar com rígidas regulamentações de segurança no campo automotivo com o objetivo de melhorar a eficiência e confiabilidade dos componentes automotivos. No entanto, uma questão desafiadora, que surge neste domínio, é lidar com uma grande quantidade de dados. Com este objetivo, neste artigo foi proposta uma infraestrutura baseada em nuvem, nomeada de manutenção de previsão automotiva na nuvem (ATOMIC do inglês, *Automotive PredicTOr Maintenance In Cloud*), para análise de prognóstico que aproveita tecnologias de Big Data e modelos matemáticos de comportamento nominal e defeituoso dos componentes automotivos para estimar on-line os indicadores de Fim de Vida (EOL do inglês, *End-Of-Life*) e RUL para os sistemas automotivos em investigação. Um estudo de caso baseado na bomba de combustível Delphi DFG1596 foi apresentado para avaliar o método de prognóstico proposto e foi realizado uma análise de benchmark das configurações de implantação da arquitetura ATOMIC em termos de escalabilidade e custo.

Medjaher, Zerhouni e Baklouti (2013) apresentaram um método de prognóstico baseado em dados utilizando técnicas de processamento de sinais e modelos de regressão. O método é aplicado em degradações aceleradas de rolamentos realizadas na plataforma experimental denominada PRONOSTIA. O objetivo do método proposto é gerar um indicador de saúde, que será utilizado para calcular o RUL. Dois sensores de aceleração são usados na plataforma PRONOSTIA para monitorar a evolução da degradação dos rolamentos testados. Os sinais de vibração relacionados aos rolamentos degradados são então comparados a um sinal de vibração nominal de um rolamento não degradado (rolamento nominal). A comparação entre os sinais é feita pelo cálculo de um coeficiente de correlação (que é considerado o indicador de saúde). Os valores do coeficiente de correlação calculado são então ajustados a um modelo de regressão que é usado para estimar o RUL.

2.4 Aquisição e processamento de dados

2.4.1 Aquisição de dados

A aquisição de dados é a primeira fase de um processo de prognóstico. Essa etapa consiste em selecionar as medições mais relevantes para a monitoração da saúde de um equipamento ou sistema. Os sensores e as estratégias de sensoriamento são fundamentais para o prognóstico. O processo de prognóstico está associado ao tipo, número e localização dos sensores, seu uso, peso, custo, faixa dinâmica entre outras características, sendo que eles podem ter ligação com fio e sem fio (VACHTSEVANOS et al., 2006). Os dados coletados pelos dispositivos transdutores raramente são usados na sua forma bruta.

A quantidade de amostras de um sinal analógico coletadas em uma determinada unidade de tempo, para conversão em um sinal digital, é denominada de taxa de amostragem. Sendo uma frequência, é comumente medida em Hertz. As amostras são os valores de um sinal analógico medidos em um determinado instante.

2.4.2 Processamento de dados

Esses dados devem ser processados apropriadamente para extrair informações úteis, resultando em uma versão reduzida dos dados originais, mas que preservam a maior quantidade de atributos destas características ou indicadores de falha, que também são indicadores de eventos de falhas e que procuram detectar, isolar e predizer a evolução ao longo do tempo. O processamento de dados consiste em realizar filtragem, compressão e correlação para remover impurezas e reduzir níveis de ruídos e volume de dados a serem processados.

Geralmente, os dados coletados possuem sinais brutos contendo ruídos e sinais irrelevantes. A eliminação desses é necessária para obter um sistema de prognóstico confiável. Essa etapa inclui a remoção de picos, a normalização de dados. Depois da obtenção de sinais relevantes e a atenuação da presença de ruído, são necessárias outras duas etapas: a extração e seleção de características.

2.4.2.1 Extração de características

Características são informações que ajudam a entender o estado de saúde de um equipamento, e que podem ser extraídas usando técnicas no domínio do tempo, frequência e tempo-frequência. As características são as propriedades heurísticas individuais mensuráveis do fenômeno em observação, sendo usualmente valores estatísticos. Dentre os extratores lineares os principais são a transformada de Fourier e a análise de componentes principais (PCA do inglês, *Principal Component Analysis*) em relação aos extratores não lineares, pode-se citar as redes neurais.

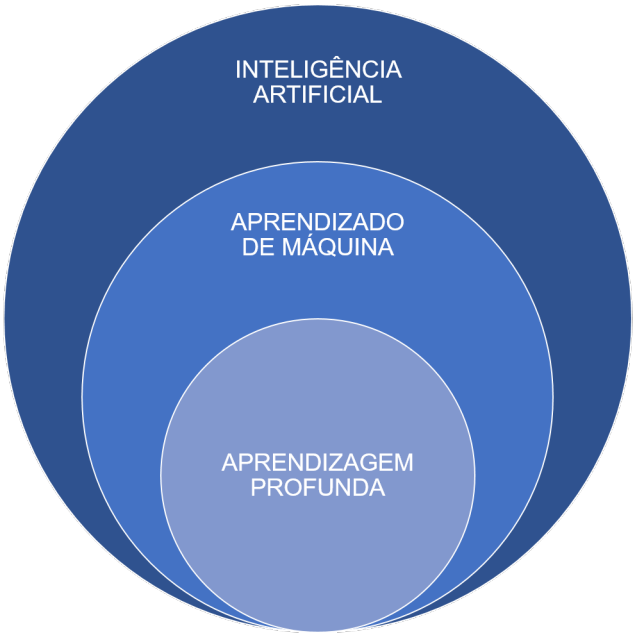
2.4.2.2 Seleção de características

Em primeira análise, extrair a maior quantidade de características é sempre melhor, já que mais características podem fornecer mais informações. Entretanto, a presença de características irrelevantes e redundantes complica o modelo de prognóstico e aumenta o custo computacional, e um fator ainda mais importante, a presença de um grande número de características, diminui a capacidade de generalização do modelo de prognóstico. O aumento do número de características pode melhorar apenas o desempenho inicial, depois de um número crítico de características, o desempenho diminui (SIMEÓN, 2015).

2.5 Aprendizagem profunda

De acordo com Oppermann (2019), a aprendizagem profunda (DL do inglês, *Deep Learning*) é um subconjunto do aprendizado de máquina (ML do inglês, *Machine Learning*), que por outro lado é um subconjunto da inteligência artificial (AI do inglês, *Artificial Intelligence*). Inteligência Artificial é um termo geral que se refere a técnicas que permitem que os computadores imitem o comportamento humano. O aprendizado de máquina representa um conjunto de algoritmos treinados em dados que tornam isso possível.

Figura 3 – Ilustração da estrutura hierárquica AI vs ML vs DL.



Fonte: Adaptada de Oppermann (2019).

A aprendizagem profunda, por outro lado, é um tipo de aprendizado de máquina, inspirado na estrutura do cérebro humano. Os algoritmos de aprendizado profundo tentam tirar conclusões semelhantes às dos humanos, analisando continuamente os dados com uma determinada estrutura lógica. Para conseguir isso, o aprendizado profundo usa uma estrutura de várias camadas de algoritmos chamadas de redes neurais artificiais (ANN do inglês *Artificial Neural Network*), na qual com base em dados históricos, consegue estimar valores futuros (OPPERMANN, 2019; REYES, 2021).

De acordo com Mathworks (2021) exemplos de aprendizagem profunda são utilizados em diversos setores, como por exemplo:

Aeroespacial e defesa: Wang et al. (2015), uma abordagem de aprendizagem profunda foi proposta e aplicada a dados de Aquisição e Reconhecimento de Alvos Móveis e Estacionários (MSTAR do inglês *Moving and Stationary Target Acquisition and Recognition*) atingindo uma precisão média de 99,1% na classificação de alvos de 10 classes, para identificar objetos de satélites.

Pesquisa médica: Urban et al. (2018) utilizaram algoritmos de aprendizado profundo para detectar automaticamente as células cancerosas, localizando e identificando pólipos, no qual se refere ao crescimento desordenado das células do tecido de um órgão, em tempo real com 96% de precisão na colonoscopia de triagem.

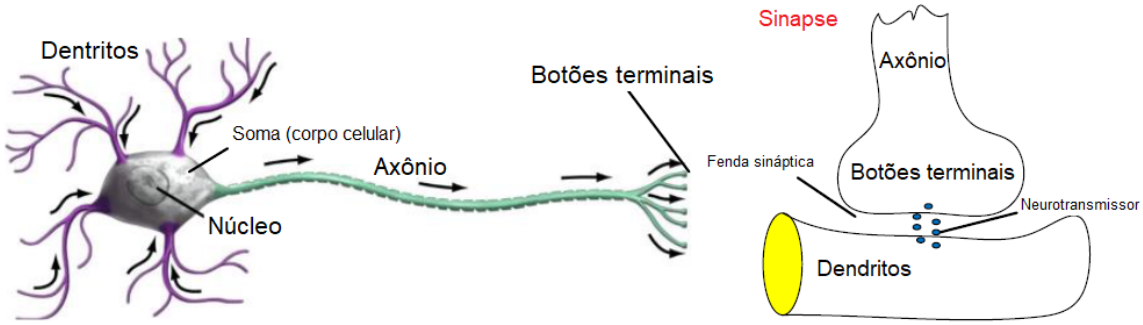
Reconhecimento de fala: Amodei et al. (2016), desenvolveram uma abordagem de aprendizado profundo que pode ser usada para reconhecer a fala da língua inglesa ou do mandarim, permitindo lidar com uma grande variedade de fala, incluindo ambientes ruidosos, sotaques e idiomas diferentes. Como resultado, em vários casos, o sistema é competitivo com a tradução humana quando comparada em conjuntos de dados padrão.

2.6 Redes neurais artificiais

De acordo com Kiranyaz et al. (2021), neurônios artificiais usados em ANN convencionais são os modelos de neurônios biológicos. No sistema nervoso dos mamíferos, o aprendizado biológico é realizado principalmente no nível celular. Conforme mostrado na Figura 4, cada neurônio é capaz de processar o sinal elétrico com base em três operações individuais:

1. Recepção das saídas dos outros neurônios por meio das conexões sinápticas nos dendritos,
2. a integração dos sinais de entrada processados no soma no núcleo da célula, e,
3. a ativação do sinal final, na primeira parte do axônio ou o chamado monte do axônio: se os potenciais agrupados excederem um certo limite, “ativa” uma série de pulsos (potenciais de ação).

Figura 4 – Ilustração de um neurônio biológico (à esquerda) com a direção do fluxo do sinal e uma sinapse (à direita).



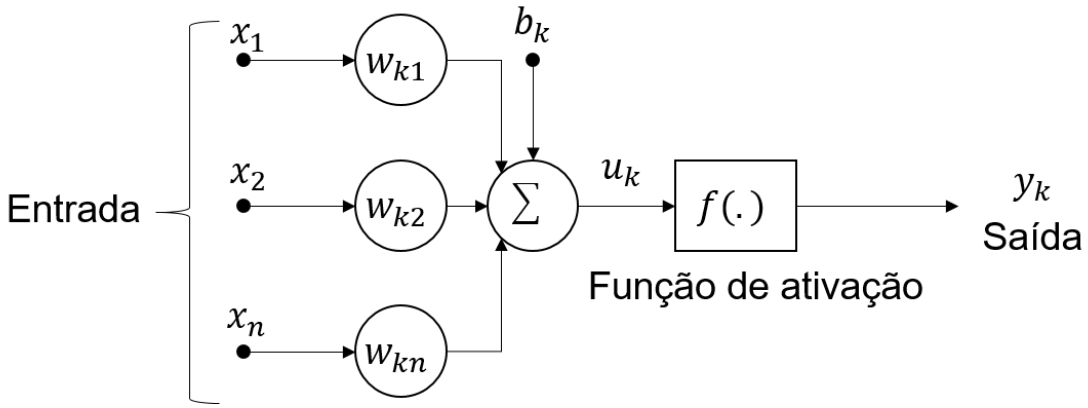
Fonte: Adaptada de Kiranyaz et al. (2021).

2.6.1 Perceptron

Segundo Kiranyaz et al. (2021), durante a década de 1940, o primeiro modelo de "neurônio artificial" foi proposto por McCulloch e Pitts (1943), que foi posteriormente usado em várias ANNs, como perceptrons multicamada (MLP do inglês *Multi Layer Perceptron*).

Conforme expresso na Equação 2.1, o neurônio artificial modela as operações básicas do neurônio biológico, onde o neurônio k com n terminais de entrada x_1, x_2, ..., x_n (dendritos), u_k é o sinal de saída do combinador linear e o sinal de saída y_k (axônio). O comportamento das sinapses é emulado através de pesos w_{k1}, w_{k2}, ..., w_{kn} acoplados as entradas dos neurônios, seguida pela adição de um termo de bias, b_k, e posteriormente a aplicação de uma função de limiar não linear, $f(.)$, que é chamada de função de ativação, descrita na Equação 2.1.

Figura 5 – Esquema de representação de um neurônio artificial, onde a soma ou corpo da célula é representada por uma composição de dois módulos, o primeiro é uma junção aditiva, somatório dos estímulos (sinais de entrada) multiplicado pelo seu fator excitatório (pesos sinápticos), e posteriormente uma função de ativação, que definirá com base nas entradas e pesos sinápticos, qual será a saída do neurônio. O axônio é aqui representado pela saída (y_k) obtida pela aplicação da função de ativação (ACADEMY, 2021).



Fonte: Elaborada pelo autor.

Matematicamente, a saída do neurônio k pode ser descrita por:

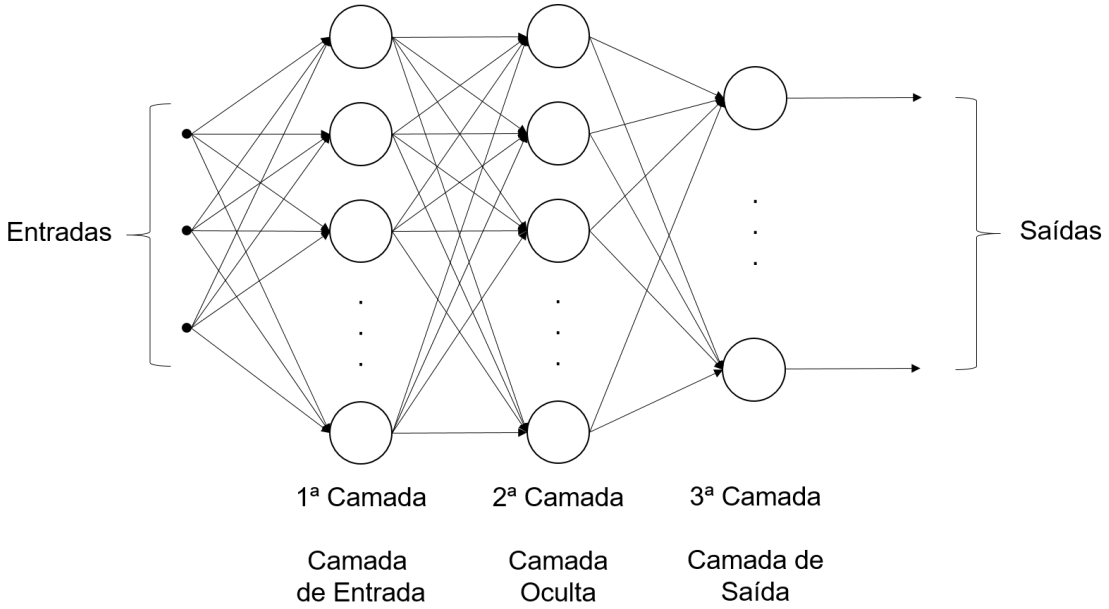
$$u_k = b_k + \sum_{j=1}^{m} w_{kj} x_j$$

$$y_k = f(u_k)$$

De acordo com Li (2019), MLP são um tipo especial de rede totalmente conectada com vários neurônios únicos. A Figura 6 ilustra o conceito de um MLP que consiste em três camadas.

O MLP representado na figura anterior tem uma camada de entrada, uma camada oculta e uma camada de saída. As unidades na camada oculta são totalmente conectadas à camada de entrada e a camada de saída está totalmente conectada à camada oculta. Pode se adicionar um número arbitrário de camadas ocultas ao MLP para criar arquiteturas de rede mais profundas.

Figura 6 – Ilustração de um MLP com três camadas.



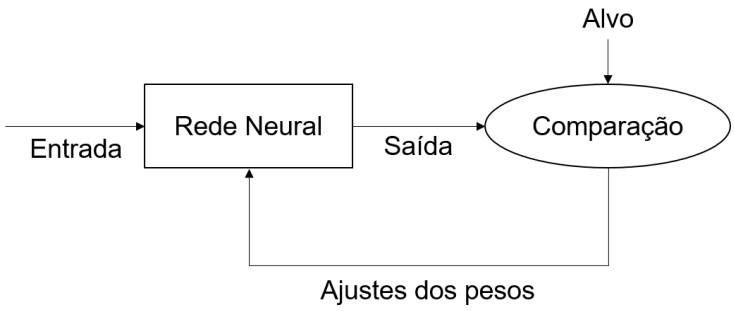
Fonte: Elaborada pelo autor.

Na prática, o número de camadas e unidades em uma rede neural são hiperparâmetros, que são variáveis que controlam o próprio processo de treinamento, que seja necessário variar para um determinado problema (REYES, 2021).

2.6.2 Treinamento supervisionado

O treinamento supervisionado necessita de um vetor de entrada e um vetor de saída desejada, conhecido como alvo. Esses dois vetores são utilizados para o treinamento da ANN. A sequência de treinamento funciona da seguinte forma: o vetor de entrada é aplicado, a saída da rede é calculada e comparada com o correspondente vetor de saída desejada, de acordo com a Figura 7. O erro encontrado então é realimentado através da rede e os pesos são atualizados pelo algoritmo de descida do gradiente, a fim de minimizar esse erro, na qual é denominado de função custo ou objetivo.

Figura 7 – Ilustração do processo de treinamento de uma ANN.



Fonte: Elaborada pelo autor.

Por exemplo, no caso da estimativa da RUL que é um problema de regressão, a função custo, E, é o erro quadrático médio (MSE do inglês, *Mean Squared Error*), sendo calculado com base na Equação 2.1:

$$e(n) = y(n) - \widehat{y}(n)$$

$$E = \frac{1}{N}\sum_{n=1}^{N}(e(n))^2$$

Onde, é o erro de previsão, $e(n)$, no n-ésimo exemplo do conjunto de dados, é a diferença entre o valor desejado, y_n, e a saída da rede, $\widehat{y}_n$.

Para treinar uma rede neural, existem duas fases: propagação direta (do inglês, *forward pass*) e retropropagação (do inglês, *backpropagation*).

Na primeira fase de propagação direta, a rede não tem nenhuma informação sobre os pesos, um método utilizado é inicializar os pesos aleatoriamente, com distribuição uniforme sobre um pequeno intervalo em torno do zero. E na sequência propagando as entradas de dados para a camada de entrada, percorrendo as camadas ocultas, e finalmente, calculando o erro da rede com base nas previsões da camada de saída (GAD, 2021).

Na retropropagação, o fluxo é revertido para que o erro propague para a camada de saída até atingir a camada de entrada passando por todas as camadas ocultas. O algoritmo de retropropagação é o conjunto de etapas usadas para atualizar os pesos da rede para reduzir o erro da rede (GAD, 2021).

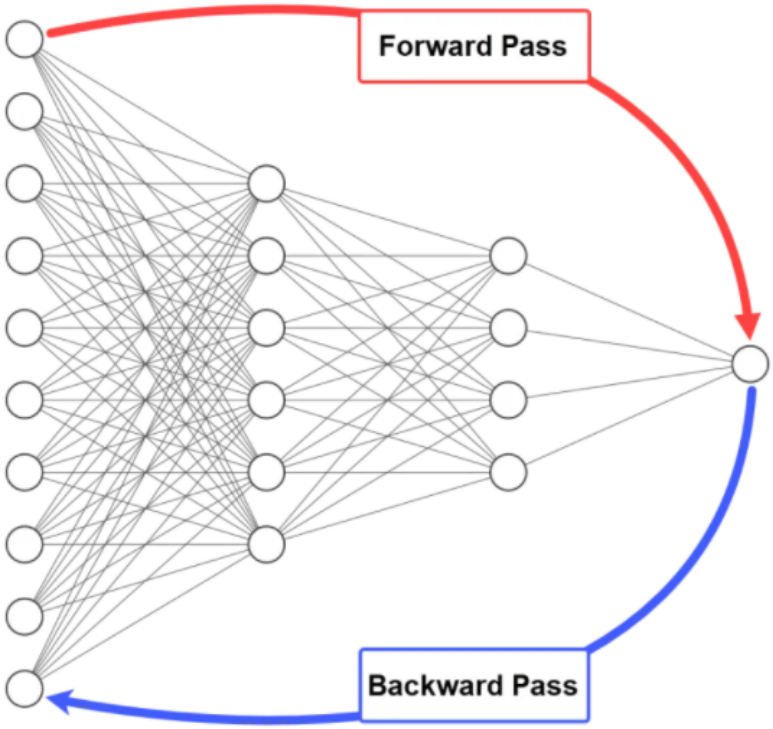
Se o erro atual for alto, a rede não aprendeu corretamente com os dados, o que pode significar que o conjunto atual de pesos não é adequado o suficiente para reduzir o erro e fazer previsões satisfatórias. Como resultado, é necessário atualizar os pesos para reduzir o erro da rede.

Segundo Gad (2021), esse algoritmo de retropropagação é genérico o suficiente para funcionar com diferentes arquiteturas de rede, como redes neurais convolucionais, redes totalmente conectadas e etc. Os únicos parâmetros do processo estão relacionados ao algoritmo de descida do gradiente, como a taxa de aprendizado.

2.6.2.1 Função de ativação

A função de ativação é a transformação não linear que causa mudança nos pesos e bias ao longo do sinal de entrada e por consequência ocasionam alterações na saída, esse fato resulta que uma rede de neurônios artificiais consiga resolver diversos problemas. Existem vários tipos de funções de ativação que executam esta tarefa de uma maneira diferente, por exemplo, *sigmoid*, *tanh* e $softmax$, De acordo com Kızrak (2019) a mais indicada para redes convolucionais é a

Figura 8 – Ilustração do processo de treinamento, onde a seta azul aponta na direção da retropropagação e a seta vermelha aponta na direção da propagação direta.



Fonte: Li (2019).

ReLu por ser mais eficiente computacionalmente e não possui grandes diferenças de acurácia quando comparada a outras funções.

Segundo Academy (2021), a unidade linear retificada (ReLU do inglês, *Rectified Linear Units*), é uma forma de função de ativação usada comumente em modelos de aprendizagem profunda. Em essência, a função retorna 0 se receber uma entrada negativa e se receber um valor positivo, a função retornará o mesmo valor positivo. A função é descrita de acordo com a Equação 2.1.

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

2.6.3 Rede Neural Convolucional

Com os avanços na neurociência cognitiva, Fukushima e Miyake (1982) propuseram o predecessora da CNN, essa foi a primeira rede que demonstrou capacidades de aprendizado compatíveis com o sistema visual biológico de um mamífero, onde os métodos baseados em CNN aprendem a extrair características diretamente dos dados brutos, Pode-se listar algumas aplicações bem sucedidas, tais como: visão computacional com Krizhevsky, Sutskever e Hinton (2012), reconhecimento de fala com Collobert e Weston (2008) e processamento de linguagem

natural com Connor, Martin e Atlas (1994).

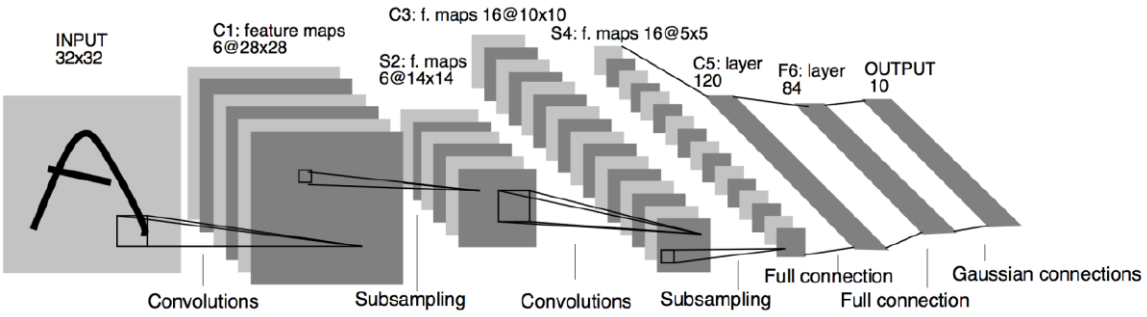
De acordo com Kiranyaz et al. (2021) as CNN tem grande potencial para identificar as padrões de sinais, sendo que cada camada pode ter um número de operadores de convolução ou agrupamento (especificados por diferentes parâmetros), de modo que vários padrões aprendidos de diferentes aspectos são considerados conjuntamente na CNN. Consequentemente, o que importa é o padrão intrínseco do sinal, em vez de suas posições ou escalas. No entanto, na estimativa RUL, é preciso considerar vários canais de séries temporais, nos quais a CNN tradicional não pode ser usada diretamente. Os desafios desse problema incluem que as unidades de processamento na CNN precisam ser aplicadas ao longo da dimensão temporal.

Para então apresentar a aplicação de CNNs, segundo Kiranyaz et al. (2021), podem ser atribuídos às seguintes vantagens:

- Podem processar sinais de entrada de tamanho variável com uma relativa eficiência computacional em comparação com as redes convencionais MLP totalmente conectadas.
- São em geral robustas a pequenas transformações nos dados de entrada, incluindo, dimensionamento, inclinação e distorção.

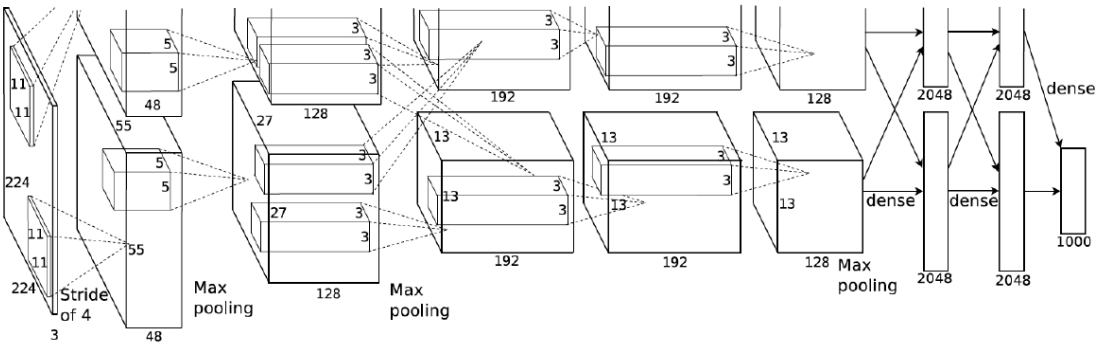
As configurações das duas CNNs mais populares, a "LeNet" (LECUN et al., 1990) e a primeira CNN profunda, "AlexNet" (KRIZHEVSKY; SUTSKEVER; HINTON, 2012) são mostradas na Figura 9 e Figura 10, respectivamente. Apesar do longo tempo entre o desenvolvimento das duas, as semelhanças conceituais e arquitetônicas são notáveis, talvez a diferença mais marcante seja a configuração "profunda" do AlexNet, que encapsula milhões de parâmetros de rede. As duas propriedades fundamentais das camadas convolucionais, "compartilhamento de peso" e "conectividade limitada" a qual consiste em várias conexões sendo controladas por um mesmo parâmetro, existem mesmo nas arquiteturas mais recentes da CNN. Na verdade, essas são as principais características que separam as CNNs das MLPs convencionais.

Figura 9 – A configuração do ancestral das CNNs, o "LeNet".



Fonte: LeCun et al. (1990).

Figura 10 – A configuração da primeira CNN profunda, a "AlexNet".



Fonte: Krizhevsky, Sutskever e Hinton (2012).

2.6.3.1 Diferenças CNN 1D x 2D

As CNNs convencionais apresentadas na seção anterior são projetadas para operar exclusivamente em dados 2D. E por isso que são frequentemente chamados de "CNN 2D". No entanto, recentemente CNNs 1D foram usadas com sucesso por Zihlmann, Perekrestenko e Tschannen (2017) para classificação de registros de eletrocardiograma, Avci et al. (2017) para detecção de danos estruturais em tempo real, Ince et al. (2016) para detecção de falha em motor em tempo real. Kiranyaz et al. (2021) evidenciou que, para certas aplicações, as CNNs 1D são vantajosas e, portanto, preferíveis aos seus correspondente 2D no tratamento de sinais 1D devido às seguintes razões:

- CNNs 1D necessitam de operações mais simples com matrizes. Isso significa que a complexidade computacional de CNNs 1D é significativamente menor do que CNNs 2D.
- Normalmente, o treinamento de CNNs 2D requerem configuração especial (por exemplo, computação em nuvem e GPUs de alto desempenho). Por outro lado, devido aos seus baixos requisitos computacionais, os CNNs 1D compactos são adequados para aplicações em tempo real e de baixo custo.

Para entender o processo de aprendizagem de uma CNN são apresentadas as definições das camadas existentes, nas arquiteturas para extração de características. Tipicamente, a arquitetura básica de uma CNN é composta de camadas alternadas de convolução e pooling, seguidas por uma ou mais camadas totalmente conectadas (LONG; SHELHAMER; DARRELL, 2014).

2.6.3.2 Camada de convolução

A convolução é uma operação matemática entre duas funções f e g, sua notação é $f * g$, esta serve para produzir uma terceira função h, que pode ser interpretada como o quanto a função f é modificada por g (RIZWAN, 2018). Ela é definida como a integral do produto de uma das funções por uma cópia deslocada e invertida da outra, a função resultante h depende do valor do

deslocamento. Se t for a variável independente e τ o deslocamento, a fórmula para a convolução continua pode ser escrita como:

$$h(t) = f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau) \cdot g(t-\tau) d\tau$$

Entretanto neste trabalho será utilizado a definição discreta da convolução, dada pelo seguinte somatório,

$$h(k) = f(k) * g(k) = \sum_{j=0}^{k} f(j) \cdot g(k-j)$$

Onde f e g são sequências de tamanho n e a fórmula fornece o valor do k-ésimo elemento do resultado.

Em redes convolucionais o primeiro argumento é a entrada (do inglês, *input*), o segundo é o filtro (do inglês, *kernel*) e a saída é denominada mapa de características (do inglês, *feature map*). No processamento de sinais, a convolução é útil para realizar filtragem ou extração de características. Mais especificamente, o uso de filtros é feito através de uma coleção de matrizes denominadas máscaras ou kernels. Essa coleção é conhecida como camada convolucional (GOODFELLOW; BENGIO; COURVILLE, 2016).

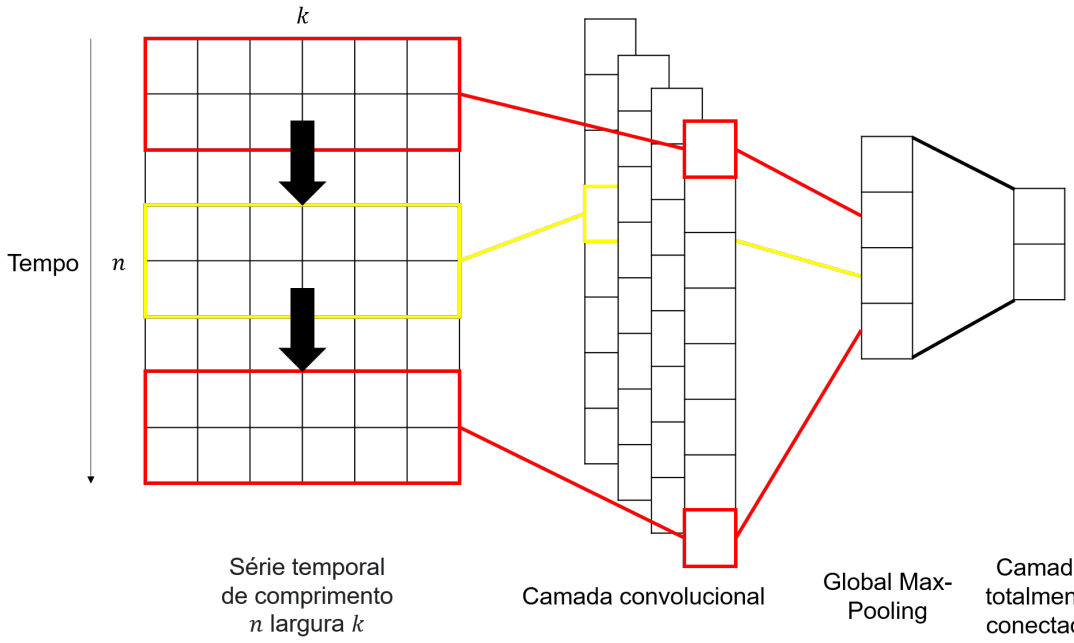
A camada de convolução é composta por diversos neurônios, cada um responsável por aplicar um filtro ao sinal. O tamanho do filtro define o tamanho da vizinhança que cada neurônio da camada irá processar. Outra variável importante para a camada convolucional é o passo (do inglês, *stride*), este valor nos diz qual será o tamanho da camada seguinte na mesma unidade. Os neurônios da mesma camada são agrupados em mapas, e estes são produzidos pelo agrupamento das saídas dos neurônios processado com um filtro em comum. Para que um conjunto de neurônios de um determinado mapa aplique o mesmo filtro em diferentes posições, é feito um compartilhamento de pesos durante o processo de treinamento. Tal compartilhamento diminui significativamente o número de parâmetros a serem aprendidos e o tempo de treinamento da rede.

Diferentemente do processo tradicional de visão computacional, no qual o modelo inicial parte da definição de filtros ou features a serem utilizadas, nas camadas de convolução das CNNs é preciso definir somente a arquitetura dos filtros: quantidade, tamanhos e stride dos filtros por camada. Com a mudança de arquitetura, as redes neurais convolucionais passam a realizar análise de campos receptivos locais (do inglês, *local receptive fields*). O processo de aprendizado da rede altera os pesos ao longo do treinamento, até encontrar os melhores valores dos filtros para o conjunto de dados utilizado (JORDAN, 2017).

Segundo Granat (2019) o uso da CNN para classificação de séries temporais (coleção de observações feitas sequencialmente ao longo do tempo) tem várias vantagens importantes em relação a outros métodos. Eles são modelos altamente resistentes a ruídos e são capazes de extrair recursos profundos e muito informativos, que são independentes do tempo.

O comprimento é o número de passos de tempo e a largura é o número de variáveis em uma série temporal multivariada. Os núcleos de convolução sempre têm a mesma largura que as séries temporais, embora sua duração possa ser variada. Dessa forma, o kernel se move em uma direção desde o início de uma série temporal em direção ao seu final, realizando a convolução. Ele não se move para a esquerda ou para a direita como acontece quando a convolução 2D usual é aplicada às imagens. De acordo com a Figura 11, onde possui uma série temporal de comprimento n e largura k, .

Figura 11 – Ilustração de uma convolução 1D para séries temporais multivariadas, os kernel são caracterizados pelos retângulos vermelhos e amarelos.



Fonte: Elaborada pelo autor.

2.6.3.3 Camada de pooling

Um aspecto chave das redes neurais convolucionais são as camadas de pooling, normalmente aplicadas após as camadas convolucionais. A ideia da camada de pooling é reduzir drasticamente o número de parâmetros. Outro importante benefício dessa camada é o aumento da invariância espacial da rede. Isso significa que a rede ficará mais robusta à possíveis torções, achatamentos ou qualquer outra variação na posição das características mapeadas (RANZATO et al., 2007).

A adição de uma camada de pooling após a camada convolucional é um padrão comum usado para ordenar camadas dentro de uma rede neural convolucional que pode ser repetida uma ou mais vezes em um determinado modelo. Especificamente, após uma não linearidade ter sido aplicada, devido a função de ativação de uma camada convolucional (BROWNLEE, 2019). Esta opera em cada mapa separadamente para criar um novo conjunto com o mesmo número de mapas agrupados.

Na camada de pooling, as características de entrada são sub-amostradas por fator adequado, de modo que a resolução dos mapas de características seja reduzida para aumentar a invariância de recursos a distorções nas entradas. Dois métodos comuns de pooling são o pooling médio e o pooling máximo, que resumem a presença média de um recurso e a presença mais ativada de um recurso, respectivamente (FUKUSHIMA; MIYAKE, 1982). Os mapas de caraterísticas de entrada são particionado pelo agrupamento médio e resultados em um conjunto de regiões não sobrepostas.

O resultado do uso de uma camada de pooling e da criação de mapas de características amostradas ou agrupadas em uma versão resumida das características detectadas na entrada. Eles são úteis porque pequenas mudanças na localização da característica na entrada detectada pela camada convolucional, terá como resultado em um mapa de características agrupada com a característica no mesmo local (BROWNLEE, 2019). Esse recurso adicionado pelo pooling é chamado de invariância do modelo.

2.6.3.4 Camada totalmente conectada

Após as camadas convolucionais e de pooling, a classificação final é feita por meio de camadas totalmente conectadas. Os neurônios em uma camada totalmente conectada têm conexões com todas as ativações na camada anterior, onde a camada totalmente conectada de uma CNN cumpre o papel de retornar uma decisão como um todo. (SANTOS et al., 2017).

Neste capítulo foi abordada a fundamentação teórica referente aos conceitos de prognóstico e CNN. No próximo capítulo será apresentado o conjuntos de dados e a metodologia de resolução do problema de estimativa da RUL.

3 METODOLOGIA

Neste capítulo apresenta-se o estudo de caso de simulação da degradação de turbinas turbofan, sendo utilizado um conjunto de dados do repositório de dados prognósticos da NASA. A simulação da degradação das turbinas foi realizada usando C-MAPSS (SAXENA et al., 2008), onde quatro conjuntos diferentes foram simulados sob diferentes combinações de condições operacionais e modos de falha e vários sensores registram a evolução da falha.

3.1 Descrição dos dados

O conjunto de dados é uma série de simulações de funcionamento até a falha. O modelo usado nas simulações foi implementado no C-MAPSS, software desenvolvido pela NASA.

3.1.1 C-MAPSS

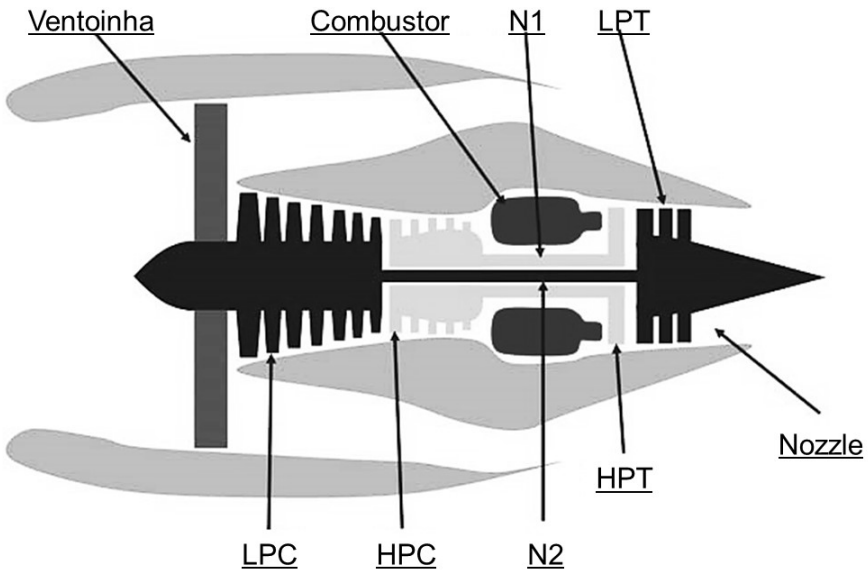
C-MAPSS é uma ferramenta para simular uma turbina comercial. O software é codificado no ambiente MATLAB® e Simulink®, e inclui um número de parâmetros de entrada editáveis que permitem ao usuário inserir valores específicos de sua própria escolha em relação ao perfil operacional, controladores de circuito fechado, condições ambientais, etc. C-MAPSS simula um modelo de turbina de classe de empuxo de 90.000 lb e o pacote inclui um modelo atmosférico capaz de simular operações em: altitudes que variam do nível do mar a 40.000 pés, número de Mach de 0 a 0,90 e temperaturas ao nível do mar de –60 a 103°F (SAXENA et al., 2008).

O pacote também inclui um sistema de gestão que permite que a turbina seja operada em uma ampla gama de níveis de empuxo em toda a gama de condições de voo, sendo possível simular defeitos em qualquer um dos cinco elementos rotativos que incluem: ventoinha, compressor de baixa pressão (LPC do inglês, *Low Pressure Compressor*), compressor de alta pressão (HPC do inglês, *High Pressure Compressor*), turbina de alta pressão (HPT do inglês, *High Pressure Turbine*) e turbina de baixa pressão (LPT do inglês, *Low Pressure Turbine*) (SAXENA et al., 2008). O diagrama da turbina na Figura 12 mostra os principais elementos do modelo da turbina e o fluxograma na Figura 13 mostra como várias sub-rotinas são montadas na simulação.

3.2 Estrutura dos dados

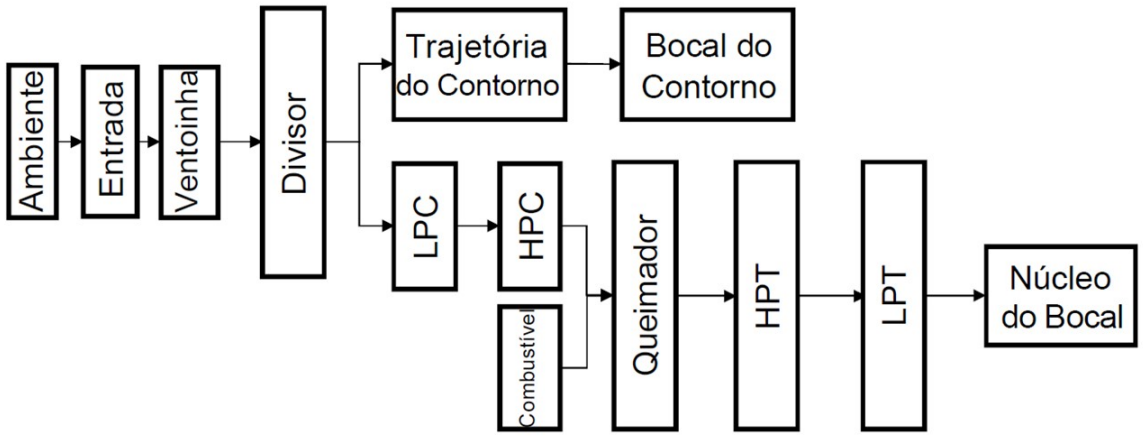
O conjunto de dados C-MAPSS é ainda dividido em 4 subconjuntos de dados, cujas as características de simulação são demonstradas na Tabela 1. Ambos os conjuntos de dados são organizados em uma matriz *n* por 26, onde *n* corresponde ao número de pontos de dados em cada componente. Cada linha é um instantâneo dos dados obtidos durante um único ciclo de tempo operacional e em 26 colunas, onde a 1ª coluna representa o número da turbina, a 2ª coluna representa o número do ciclo operacional, da 3ª até a 5ª coluna representa as três configurações

Figura 12 – Diagrama simplificado da turbina simulada no C-MAPSS.



Fonte: Adaptada de (SAXENA et al., 2008)].

Figura 13 – Layout mostrando vários módulos e suas conexões conforme modelado na simulação



Fonte: Adaptada de (SAXENA et al., 2008)

operacionais e da 6ª até a 26ª coluna representa os 21 valores dos sensores. Mais informações sobre os 21 sensores podem ser encontradas em Saxena et al. (2008).

O desempenho da turbina pode ser afetado significativamente por três configurações operacionais nos dados (BABU; ZHAO; LI, 2016). Cada trajetória dentro das trajetórias de treino e teste é considerada o ciclo de vida de uma turbina. Embora cada turbina seja simulada com

diferentes condições iniciais, essas condições são consideradas normais (sem falhas). Para cada trajetória da turbina dentro dos conjuntos de treinamento, a última entrada de dados corresponde ao momento em que a turbina é declarada insalubre ou com status de falha. Por outro lado, os conjuntos de teste contêm dados algum tempo antes da falha. Para cada um dos conjuntos de dados C-MAPSS, o valor RUL real das trajetórias de teste foram disponibilizadas ao público (BABU; ZHAO; LI, 2016).

Esses conjuntos de dados possuem características únicas que os tornam muito úteis e adequados para o desenvolvimento de algoritmos de prognóstico (RAMASSO; SAXENA, 2014).

1. Os dados representam uma resposta multidimensional de um sistema não linear complexo de uma simulação de alta fidelidade que modela muito de perto um sistema real.

2. Essas simulações incorporaram altos níveis de ruído introduzidos em vários estágios para acomodar a natureza da variabilidade geralmente encontrada.

3. Os efeitos das falhas são mascarados devido às condições operacionais, que é outra característica comum da maioria dos sistemas.

4. Dados de várias unidades são fornecidos para permitir que algoritmos extraiam tendências e construam associações para aprender o comportamento do sistema útil para prever RUL.

Tabela 1 – Detalhes dos conjuntos de dados (Simulado no C-MAPSS)

Conjunto de dados	Subconjuntos	Modos de falha	Modos de operação	Turbinas para treinamento	Turbinas para teste
Simulação de degradação de turbina	FD001	1	1	100	100
	FD002	1	6	260	259
	FD003	2	1	100	100
	FD004	2	6	249	248

Fonte: Adaptada de (SAXENA et al., 2008)

3.3 Seleção do subconjunto de dados

Para desenvolver este trabalho, utilizou-se o conjunto FD001, pois de acordo com a pesquisa de Ramasso e Saxena (2014) este subconjunto foi utilizado em 70% das publicações em comparação com todos os outros subconjuntos de dados entre os quatro do repositório da NASA. Isso pode ser atribuído à sua aparente simplicidade em comparação com o resto, porque algumas das medições dos sensores neste conjunto representam uma tendência ordenada. A Tabela 2 apresenta a estrutura do conjunto de dados.

A Tabela 3 apresenta a descrição das grandezas monitoradas, detalhando o número do sensor, etiqueta, descrição da variável e a grandezas físicas.

Tabela 2 – Exemplo da estrutura dos dados FD001.

Número da turbina	Ciclo operacional	Configuração operacional 1	...	Configuração operacional 3	Sensor 1	...	Sensor 21
1	1	0,0070	...	100	518,67	...	23,4190
1	2	0,0019	...	100	518,67	...	23,4236
:	:	:	:	:	:	:	:
1	100	0,0009	...	100	518,67	...	22,9649

Fonte: Adaptada de Simeón (2015)

Tabela 3 – Saídas do C-MAPSS para medir a resposta do sistema.

Sensor	Etiqueta	Descrição	Grandeza
1	T2	Temperatura total na entrada da ventoinha	°R
2	T24	Temperatura total na saída do LPC	°R
3	T30	Temperatura total na saída do HPC	°R
4	T50	Temperatura total na saída do LPT	°R
5	P2	Pressão na entrada da ventoinha	psia
6	P15	Pressão total no duto by-pass	psia
7	P30	Pressão total na saída do HPC	psia
8	Nf	Velocidade da ventoinha	rpm
9	Nc	Velocidade do núcleo	rpm
10	epr	Taxa de pressão da turbina (P50/P2)	–
11	Ps30	Pressão estática na saída do HPC	–
12	phi	Taxa de fluxo de combustível	–
13	NRf	Velocidade corrigida do ventilador	rpm
14	NRc	Velocidade corrigida do núcleo	rpm
15	BPR	Taxa de desvio	–
16	farB	Taxa de queima combustível-ar	–
17	htBleed	Entalpia de dreno	–
18	Nf_dmd	Velocidade exigida da ventoinha	rpm
19	PCNfR_dmd	Velocidade exigida corrigida da ventoinha	rpm
20	W31	Dreno do refrigerante HPT	lbm/s
21	W32	Dreno de refrigerante LPT	lbm/s

Fonte: Adaptada de Simeón (2015)

Os dados do arquivo “train_FD001.txt” foram utilizados para treinamento e os dados do arquivo “test_FD001.txt” para validação, ambos os arquivos contém informações de 100 turbinas, consideradas do mesmo tipo e cada ciclo de operação está contaminado com ruído (RAMASSO; SAXENA, 2014).

Foram utilizados dados completos treinamento/teste para validar a abordagem desenvolvida. Os valores reais de RUL foi adquirido do arquivo “rul_FD001.txt” para comparação com os valores de RUL estimados por uma CNN 1D.

3.4 Pré-processamento dos dados

3.4.1 Seleção das variáveis

Para a realização do prognóstico, a seleção das variáveis é baseada na tendência, que é o comportamento ao longo da série temporal, que pode ser causada pela mudança gradual da condição das turbinas, ou qualquer outro aspecto que afete a variável de interesse no longo prazo que pode aumentar ou diminuir em função dos ciclos operacionais. Além disso, a seleção de variáveis reduz a dimensionalidade do problema, elimina atributos irrelevantes ou redundantes, portanto, melhora a eficácia dos classificadores (LIU; GEBRAEEL; SHI, 2013).

Alguns sensores não mostram uma tendência clara como outros devido ao alto ruído ou sua baixa sensibilidade à degradação. Incluí-los na análise pode diminuir a precisão da previsão. As medições descartadas foram aquelas com valores constantes ao longo de todo o ciclo de vida útil das turbinas (ISHIBASHI; JÚNIOR, 2013; GOURIVEAU; ZERHOUNI, 2012). Do total de variáveis que contém o conjunto de dados, dentre as 21 variáveis monitoradas, apenas os sensores com uma tendência são selecionados para processamento.

A fim de avaliar se todos os sensores selecionados ajudam a melhorar a precisão da estimativa de RUL, ou seja, se o uso de um subconjunto desses sensores selecionados pode realmente melhorar o desempenho, os procedimentos de seleção das variáveis a seguir deste trabalho, foram baseados nos principais conjuntos da literatura.

Conjunto 01: Este conjunto leva em consideração todas as medições dos sensores que apresentam qualquer variação temporal, sendo este o mais comumente utilizado, pois não é necessário nenhuma abordagem de extração de recursos.

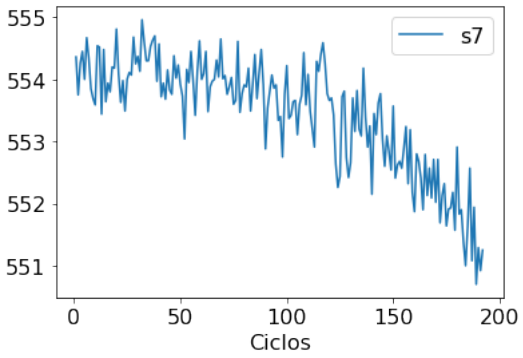
Conjunto 02: Os estudos experimentais de Wang et al. (2008) nos conjuntos de dados de teste mostraram que a escolha de sete sensores produziu uma melhor estimativa geral de RUL, sendo eles s2, s3, s4, s7, s11, s12, s15.

Conjunto 03: Em Wang et al. (2008), os autores propuseram usar medições de sensores s2, s3, s4, s7, s11, s12 e s15, mas de acordo com Ramasso, Rombaut e Zerhouni (2012) o s12 deve ser substituído por s13 porque s12 é quase constante e não traz informação, sendo eles s2, s3, s4, s7, s11, s13, s15.

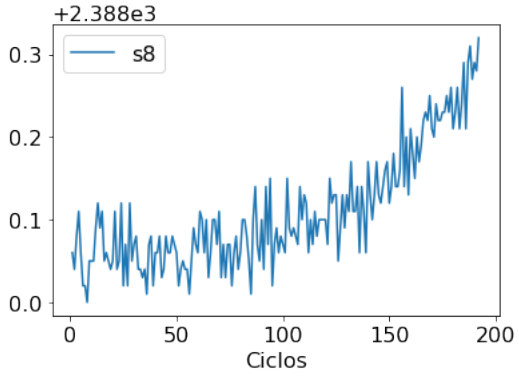
Conjunto 04: Complementar às técnicas tradicionais baseadas em modelos para prognósticos de falhas, Sarkar, Jin e Ray (2011) propõem filtragem dinâmica particionando de forma otimizada os dados da série temporal de observação do sensor sendo o objetivo mascarar os efeitos da variação do nível de ruído do sensor e ampliar as assinaturas de falha do sistema, sendo eles s7, s8, s12, s16, s17 e s20.

Conjunto 05: Este conjunto de sensores foi baseado a partir da pesquisa realizada por Simeón (2015), no qual recomenda a utilização dos sensores s7, s8, s9, s12, s16, s17 e s20. De acordo com a Figura 14.

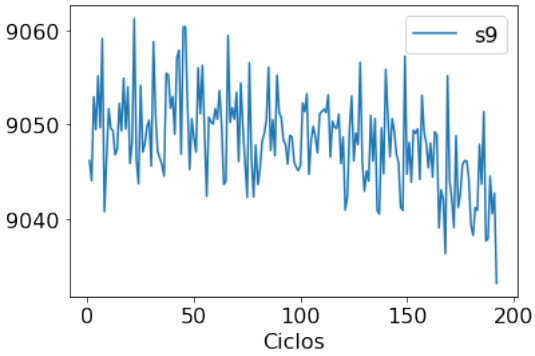
(a) P30: Pressão na saída do HPC [psia]



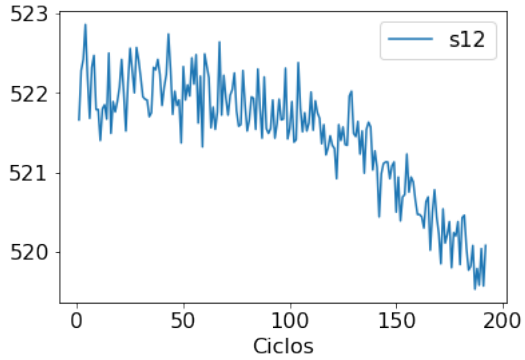
(b) Nf: Velocidade da ventoinha [rpm]



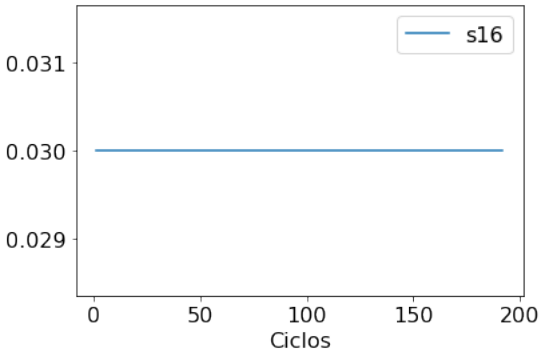
(c) Nc: Velocidade do núcleo [rpm]



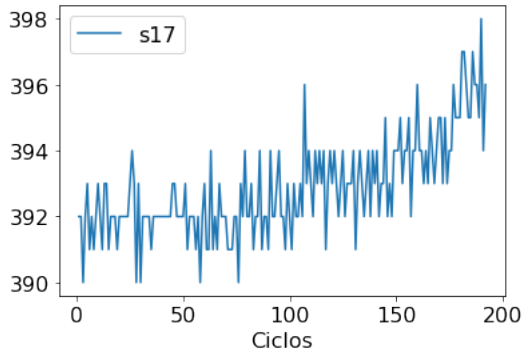
(d) phi: Taxa de fluxo de combustível



(e) farB: Taxa de queima combustível-ar



(f) htBleed: Entalpia de dreno



(g) W31: Dreno do refrigerante HPT [lbm/s]

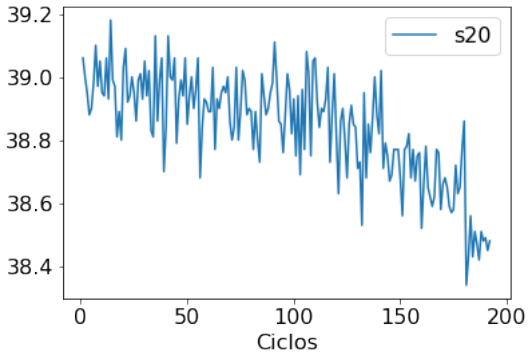


Figura 14 – Ilustração das variáveis selecionadas do conjunto 05 na turbina 1.

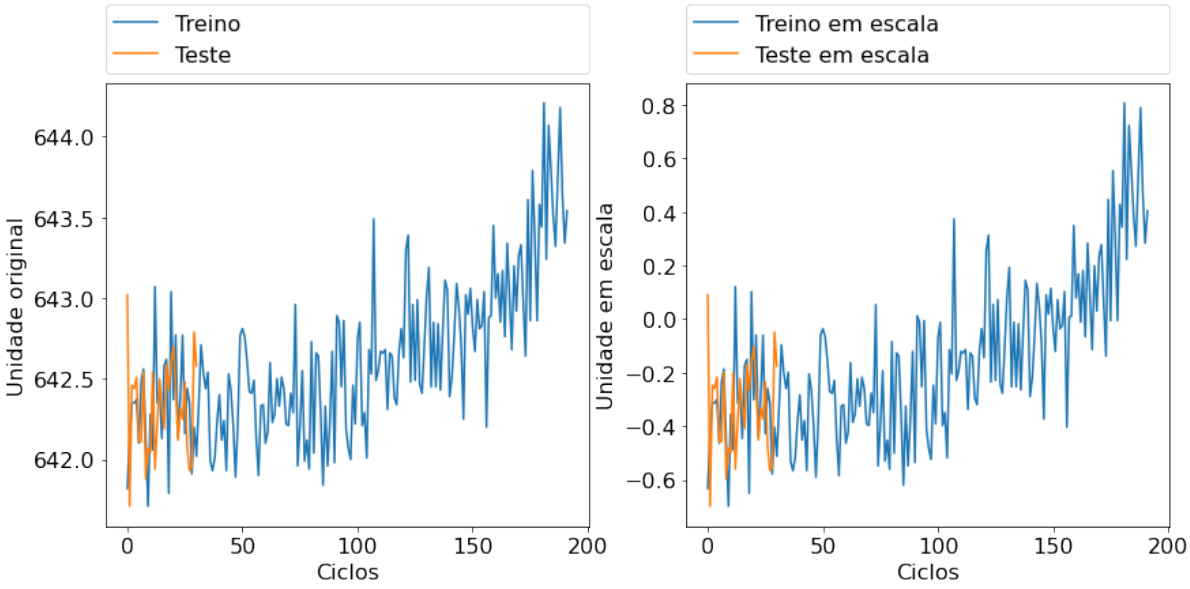
3.4.2 Normalização

Antes de realizar qualquer treinamento e teste, foi feito a normalização dos dados para que os pontos fiquem dentro de uma faixa razoável de valores, conforme a Equação 3.1.

$$z = \frac{x - \overline{x}}{\sigma} \tag{3.1}$$

Onde, $\overline{x}$ é a média da amostra, σ é o desvio padrão da amostra, o valor absoluto de z representa a distância do valor x entre a média da amostra em unidades do desvio padrão.

Figura 15 – Medições do sensor 2 da turbina 1 normalizado.



Fonte: Elaborada pelo autor.

3.4.3 Janela deslizante

Para realizar a análise dos dados, é utilizada a técnica da janela deslizante retangular de tamanho fixo, no qual a ideia principal é segmentar o sinal da série temporal em uma coleção de pedaços de sinais, sendo usada uma matriz de tamanho constante, que está se movendo da esquerda para a direita em função dos sensores, de acordo com a Figura 16.

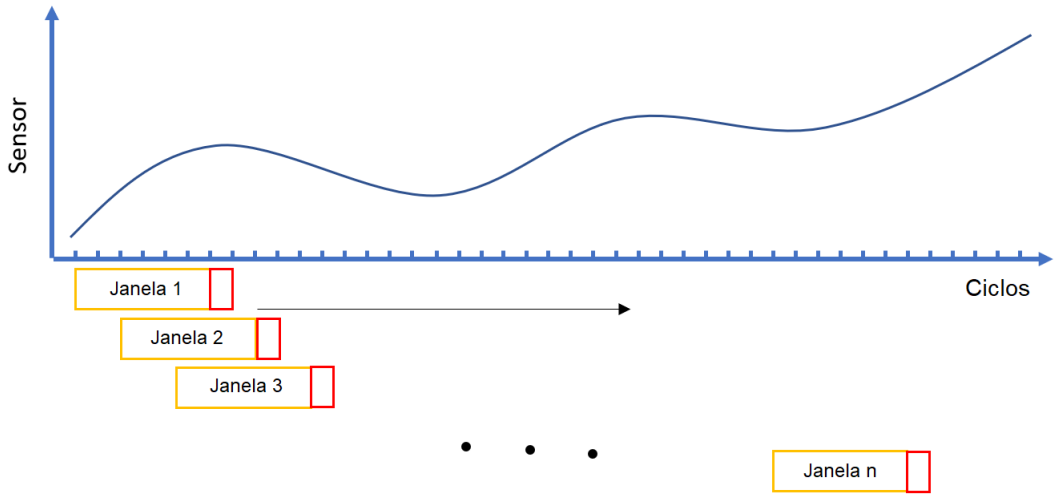
Na etapa de treinamento da CNN, esses pedaços de sinais servem de dados de entrada na rede, sendo divididos em séries temporais com uma janela fixa de 30 ciclos e um deslizamento de 1 passo.

Por exemplo, a turbina 1 tem 192 ciclos de treino, com um comprimento de janela igual a 30 extraímos 162 séries com comprimento 30:

- Janela 1 ⇢ do ciclo 1 ao ciclo 31,
- Janela 2 ⇢ do ciclo 2 ao ciclo 32,

- ... ,
- Janela n $\dashrightarrow$ do ciclo n para ciclo $n+30$.

Figura 16 – Janela deslizante para um único sensor.



Fonte: Elaborada pelo autor.

Os valores da janela retangular fixa de 30 ciclos e o deslizamento de 1 passo foram obtidos empiricamente, sempre visando minimizar o RMSE e a Pontuação S, descritas na Seção 3.6.

3.4.4 Função alvo RUL

Em sua forma mais simples, os algoritmos de prognóstico são semelhantes aos problemas de regressão. No entanto, ao contrário dos problemas de regressão típicos, um desafio inerente para problemas de prognóstico baseados em dados é determinar os valores de saída desejados para cada ponto de dados de entrada.

Neste trabalho foi proposta a comparação de duas funções alvo RUL: RUL linear e RUL linear por partes.

3.4.4.1 RUL linear

Esta abordagem é a solução usualmente mais utilizada devido simplesmente atribuir a saída desejada como o tempo real restante antes da falha funcional, de acordo com a Figura 17a. Essa abordagem, entretanto, implica inadvertidamente que a integridade do sistema se degrada linearmente com o uso. Deve-se notar que nos casos em que o conhecimento de um modelo de degradação adequado não está disponível, o modelo linear é a escolha mais natural para uso.

(a) RUL linear **(b) RUL linear por partes**

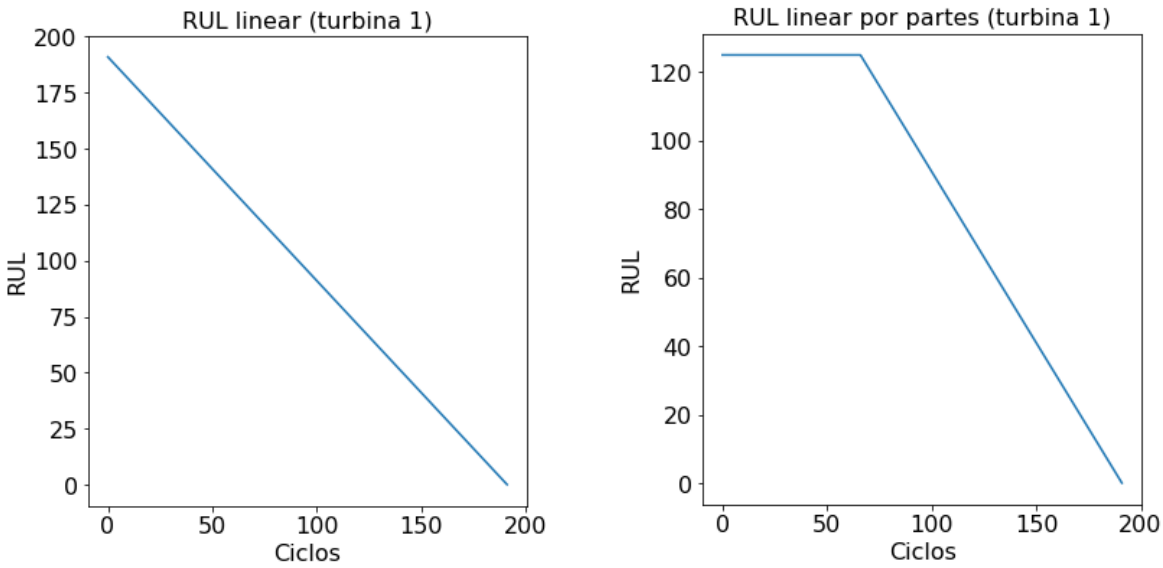


Figura 17 – Funções alvo RUL.
Fonte: Elaborada pelo autor.

3.4.4.2 RUL linear por partes

Esta abordagem alternativa deriva os valores de saída desejados com base em um modelo de degradação linear por partes. Este um modelo de degradação foi proposto em Hou, Pi e Li (2020), Elsheikh, Yacout e Ouali (2019) e Li, Ding e Sun (2018), limitando o valor máximo da função RUL em 125, de acordo com Figura 17b.

Neste modelo, no início, RUL é atribuído a um número fixo. O número fixo é chamado de RUL inicial de 125. Quando o valor dos ciclos atinge 75, a RUL segue um modelo de degradação linear.

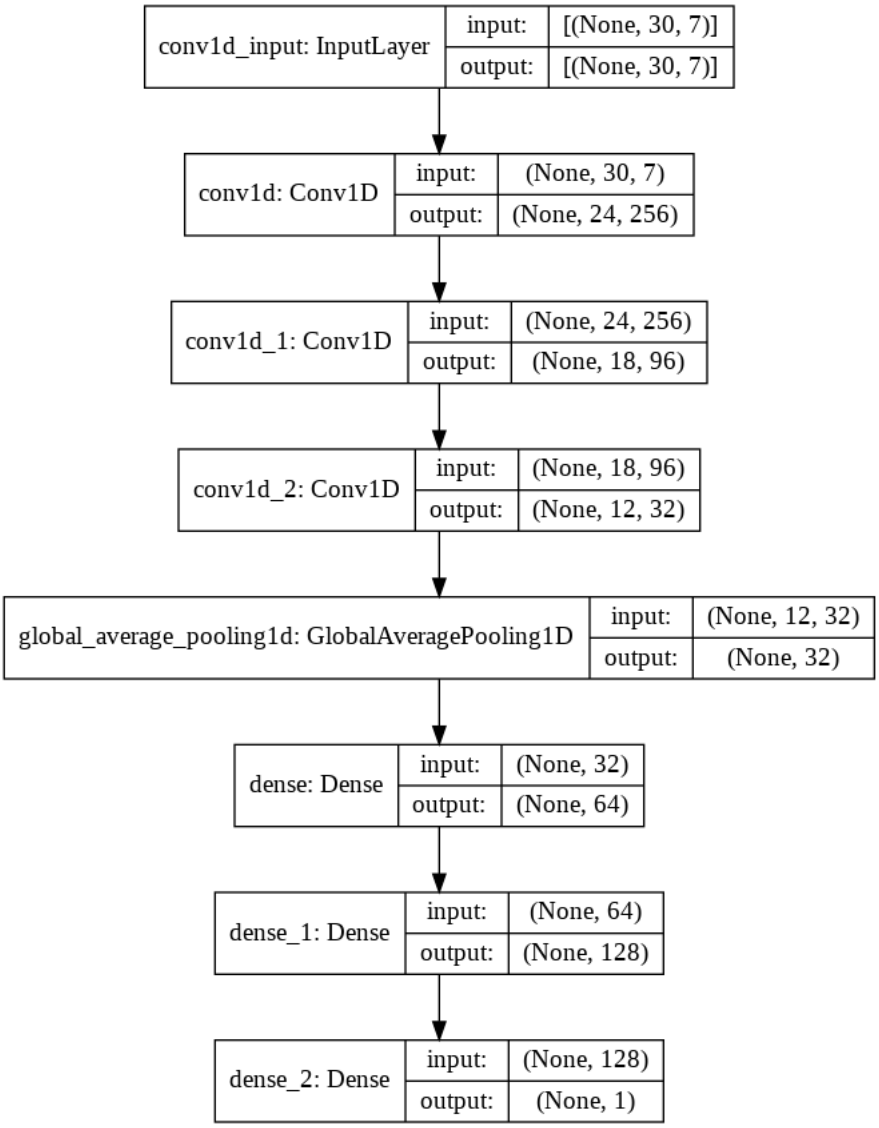
Segundo Elsheikh, Yacout e Ouali (2019) a função alvo RUL linear por partes tem maior probabilidade de evitar que o algoritmo superestime o RUL. Além disso, também é um modelo mais lógico, pois a degradação do sistema normalmente só começa após um certo grau de uso.

3.5 Arquitetura da CNN

A arquitetura da CNN foi baseada principalmente nos trabalhos desenvolvidos por Krizhevsky, Sutskever e Hinton (2012) e LeCun et al. (1990).

Nesta arquitetura proposta, conforme mostrado na Figura 18, são utilizadas três camadas de convolução, com todas apresentando uma função de ativação ReLu na sua saída, em seguida uma camada de pooling que realiza a operação de agrupamento médio global, e por fim apresenta três camadas totalmente conectadas para estimar o valor da RUL.

Figura 18 – Proposta de arquitetura CNN 1D para estimativa da RUL.



Fonte: Elaborada pelo autor.

3.6 Métricas de desempenho

3.6.1 Pontuação S

A função de pontuação usada neste trabalho é idêntica àquela usado no PHM 2008 Data Challenge, uma competição aberta a todos os participantes em potencial. O objetivo da conferência era estimar a RUL de um componente não especificado usando técnicas baseadas em dados, sendo o erro entre o valor estimado e o valor verdadeiro da RUL, calculado de acordo com a Equação 3.2 . Essa função de pontuação é ilustrada na Equação 3.3, onde, i é a última estimativa da rede para cada turbina, N é o número de turbinas no conjunto de teste, S é a

pontuação calculada:

$$h_i = \text{RUL}_{estimado} - \text{RUL}_{verdadeiro} \tag{3.2}$$

$$\begin{aligned} s_i &= \begin{cases} e^{-\frac{h_i}{13}} - 1, & h_i < 0 \\ e^{-\frac{h_i}{10}} - 1, & h_i \geq 0 \end{cases} \\ S &= \sum_{i=1}^{N} s_i \end{aligned} \tag{3.3}$$

Esta função de pontuação penaliza as previsões tardias (tarde demais para realizar a manutenção) mais do que as previsões iniciais (sem grandes danos, embora possa desperdiçar recursos de manutenção). Isso está de acordo com a atitude adversa ao risco nas indústrias aeroespaciais. No entanto, existem várias desvantagens com essa função. A desvantagem mais significativa sendo um único outlier (com uma previsão muito tardia) dominaria a pontuação de desempenho geral, mascarando assim a verdadeira precisão geral do algoritmo.

Outra desvantagem é a falta de consideração do horizonte prognóstico do algoritmo. O horizonte de prognóstico avalia o tempo antes da falha em que o algoritmo é capaz de estimar com precisão o valor RUL dentro de um certo nível de confiança. Finalmente, esta função de pontuação favorece algoritmos que reduzem artificialmente a pontuação subestimando o RUL. Apesar de todas essas deficiências, a função de pontuação ainda é usada neste trabalho para fornecer resultados de comparação com outros métodos na literatura.

3.6.2 Raiz do erro quadrático médio

Além da função de Pontuação S, o RMSE de RUL estimados também é empregado como uma medida de desempenho. O RMSE é escolhido porque dá peso igual às previsões iniciais e tardias. Usar RMSE em conjunto com a função de pontuação evitaria favorecer um algoritmo que diminui artificialmente a pontuação ao subestimá-la, mas resultando em RMSE mais alto. O RMSE é definido conforme Equação 3.4 a seguir:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} h_i^2} \tag{3.4}$$

Pode-se observar que em valores de erro absoluto mais baixos a função de pontuação resulta em valores mais baixos do que o RMSE. As características relativas das métricas de avaliação do desempenho serão úteis durante a discussão dos resultados experimentais na parte posterior deste trabalho.

4 RESULTADOS E DISCUSSÕES

Esta seção descreve o algoritmo usado, os parâmetros de configuração, a topologia, e os resultados de prognóstico obtidos por uma CNN unidimensional. Os algoritmos foram desenvolvidos em Python baseados no Keras sendo uma Interface de Programação de Aplicação (API do inglês, *Application Programming Interface*) de alto nível do TensorFlow para criar e treinar o modelo de aprendizado profundo. A etapa de treinamento é realizada usando 80% dos dados do arquivo "train_FD001.txt" e 20% para validação do conjunto de dados do arquivo "train_FD001.txt". Na etapa de teste é utilizado o conjunto de dados do arquivo "test_FD001.txt" obtendo na saída a RUL estimada. Os valores de RUL obtidos na saída são comparados com os valores de RUL desejados que se encontram no arquivo "rul_FD001.txt".

4.1 CNN 1D

O prognóstico de RUL do conjunto de turbinas é realizado através de uma CNN 1D. Considera-se como entradas, as grandezas dos sensores, e como saída uma variável contínua que representa o número de ciclos restantes até a ocorrência de uma falha. Esta variável representa a RUL e é considerada como a saída desejada durante o treinamento da CNN 1D. Considerando que o final da sequência de dados representa o último ciclo antes da ocorrência da falha de uma turbina específica.

Os resultados obtidos com a CNN 1D serão utilizados como referência de comparação com outros métodos encontrados na literatura. O experimento da definição de parâmetros de treinamento e teste da CNN 1D é controlado pelo Algoritmo 1, baseado em Sahoo (2018).

Foram treinadas diferentes redes para testar os conjuntos de sensores baseados na literatura técnica e mantendo fixo o restante dos parâmetros, avaliando o desempenho da CNN 1D para cada experimento, utilizando o tamanho da janela de tempo de 30 para FD001.

4.2 Avaliação de desempenho

A avaliação de desempenho do subconjunto FD001 é realizada comparando as RUL estimadas com as RUL desejadas fornecidas no arquivo de texto "rul_FD001.txt". Através das equações definidas nas Eq. 3.3 e Eq. 3.1, obtém-se os valores das métricas de prognóstico utilizadas para realizar uma comparação quantitativa.

Além disso, estas abordagens também enfrentam outras desvantagens da escolha humana de parâmetros que podem limitar sua aplicabilidade. Trabalhos dedicados a análise de agrupamento séries temporais são relativamente escassos para prognósticos. Certamente é difícil alcançar uma estimativa apurada de RUL devido a problemas relacionados aos dados reais de

Algoritmo 1: Pseudo-código do Experimento CNN 1D

Entrada: Dados de Treinamento, Dados de Teste, RUL real
Saída: RUL estimada, métricas de prognóstico
01: **Início**
02: Carregar o conjunto de dados de Treinamento;
03: Carregar o conjunto de dados de Teste;
04: Carregar o conjunto de dados de RUL real;
05: Pré-processamento dos dados
06: Seleção das sensores
07: Normalização
08: Janela deslizante
09: Função alvo RUL
10: Época = 1;
11: **Repita**
12: Definir Parâmetros CNN 1D;
13: Gerar Pesos aleatórios;
14: Treinar a CNN 1D;
15: Testar a CNN 1D (Estimar a Vida Útil Remanescente);
16: Analisar desempenho (Métricas de Prognóstico);
17: **Até** (Época $\leq$ 30 Épocas)
17: Salvar os Melhores Parâmetros, Pesos, RUL estimada e as métricas de desempenho;
18: **Fim**

monitoramento, ou seja, ruído do sensor e alta variabilidade, ausência de conhecimento prévio, dimensão e escalas, bem como tamanho dos dados, etc.

Então a fim de avaliar o desempenho da estimativa RUL foram utilizados diferentes conjunto de sensores envolvidos e dois métodos de degradação.

4.2.1 CNN 1D com RUL linear

Essa abordagem considera o método de degradação a função alvo RUL linear, demonstrada na Seção 3.4.4 para estimar a RUL das turbinas. Para realizar uma análise sistemática o conjunto de sensores foi variado de acordo com as principais referências técnicas.

Nas Tabelas 4, 5 e 6, os valores estatísticos e as métricas de desempenho são calculados apenas a partir do erro da última estimativa da rede para cada turbina de acordo com a Equação 3.2. De modo mais específico são 100 turbinas de teste, então a CNN estima 100 valores de RUL (RUL estimado) um valor para cada turbina. E a saída desejada é obtida através dos dados do arquivo "rul_FD001"(RUL verdadeira), disponibilizada no conjunto de dados. A partir desses valores é calculado o erro, e seus respectivos valores estatísticos e métricas de desempenho.

Para interpretar as Tabelas 4, 5 e 6, é necessário ter o entendimento que o erro calculado é dado em ciclos. Na segunda coluna apresenta o valor médio do erro calculado, $\overline{h}$, então o melhor

cenário é um valor absoluto mais próximo possível de zero. A terceira coluna apresenta o valor do desvio padrão do erro, σ, quando o desvio é baixo quer dizer que os erros estão próximos da média. A terceira e quarta coluna apresentam os valores máximos e mínimos dos 100 erros calculados. A quinta coluna é a métrica responsável por avaliar o desempenho da estimação da rede, lembrando que o RMSE quanto mais próximo de zero for, melhor o resultado para as competições de prognóstico. Por fim, a sexta coluna é o percentual de amostras acima de zero,

Tabela 4 – Resultados da CNN 1D com RUL linear

	$\overline{h}$	σ	Min	Máx	RMSE	Acima
Conjunto 01	14,49	41,76	-41,82	110,57	30,67	70%
Conjunto 02	**11,51**	27,34	-49,64	93,42	29,00	67%
Conjunto 03	13,38	26,80	-54,19	**74,38**	29,24	**65%**
Conjunto 04	23,38	**23,38**	**-30,35**	109,66	**28,05**	72%
Conjunto 05	17,87	24,78	-39,29	101,95	29,62	71%

Fonte: Elaborada pelo autor.

Observando os resultados da Tabela 4 nota-se que os conjuntos 02 e 03 apresentam em termos de média os menores valores. Segundo esta medida estatística pode-se dizer que os conjuntos 02 e 03, em média possuem um erro inferior quando comparada com os outros conjuntos. Por outro lado estes mesmos conjuntos apresentam valores intermediários de desvio padrão, o que pode se tornar um problema no planejamento da manutenção, pois quanto maior o desvio padrão, maior será a dispersão e mais afastados da média estarão os erros calculados, de forma prática esses valores podem induzir uma manutenção precoce ou tardia, o que vem acarretar em custos para o setor responsável. Comparando o desempenho através da métrica RMSE, todos os conjuntos, apresentam um valor muito próximo.

Depois da análise comparativa, em termos gerais verificou-se que os resultados do conjuntos 02 e 03 são equivalentes e satisfatórios, pois os mesmos apresentam valores estatísticos intermediários, média e desvio padrão, e apresentam os menores valores de erros acima de 0, em torno de 67% e 65%, respectivamente, resultando em conjuntos de sensores aceitáveis.

Todos os resultados obtidos para RUL linear são demonstrados na Figura 19, na qual apresenta o histograma dos erros e apresenta uma linha vertical tracejada em vermelho em zero, para servir de referência.

4.2.2 CNN 1D com RUL linear por partes

Essa abordagem considera o método de degradação a função alvo RUL linear por partes, demonstrada na Seção 3.4.4 para estimar a vida útil restante das turbinas. Para realizar uma análise sistemática o conjunto de sensores foi variado de acordo com as principais referências técnicas.

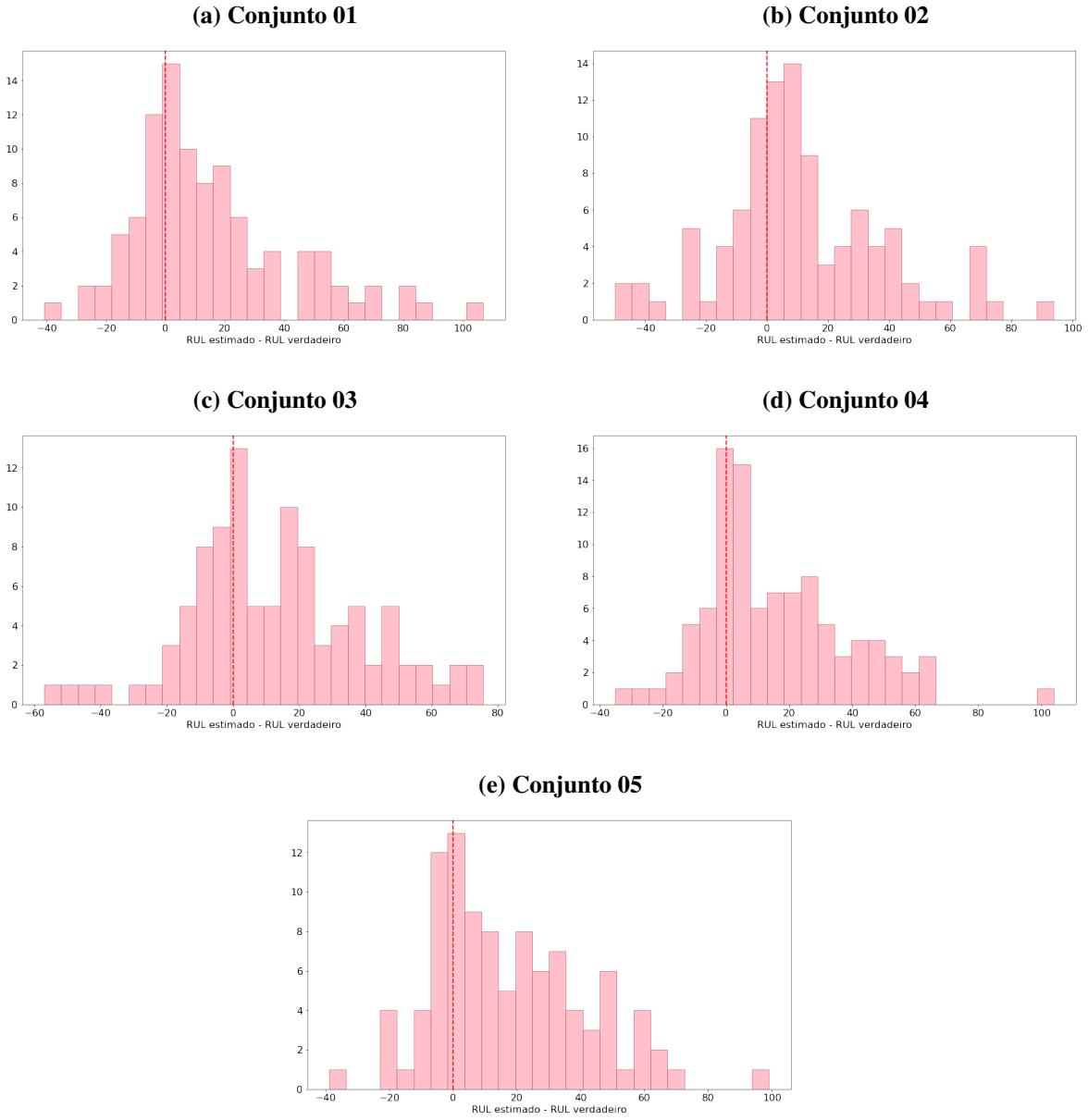


Figura 19 – Distribuição dos erros das estimativas da RUL para CNN 1D com RUL linear

Observando os resultados da Tabela 5 nota-se que o conjunto 05 apresenta em termos de média o menor valor absoluto. Segundo esta medida estatística pode-se dizer que o conjunto 05, em média tem um erro inferior quando comparada com os outros conjuntos. E também este mesmo conjunto apresenta um dos menores valores de desvio padrão, o que pode se tornar um fator benéfico no planejamento da manutenção, pois quanto menor o desvio padrão, menor a dispersão em torno da média estarão os erros calculados. Comparando o desempenho dos conjuntos através da métrica RMSE, resulta que todos apresentam um valor muito próximo.

Depois da análise comparativa, em termos gerais verificou-se que os resultados do conjunto 05 são satisfatórios, pois o mesmo apresenta os menores valores estatísticos, média e desvio padrão, e apresenta o valor de amostras acima de 0, em torno de 47%, resultando em um

Tabela 5 – Resultados da CNN 1D com RUL linear por partes

	$\overline{h}$	σ	Min	Máx	RMSE	Acima
Conjunto 01	0,84	**16,12**	**-38,60**	40,71	**16,06**	58%
Conjunto 02	-0,59	17,49	-45,65	45,59	17,41	52%
Conjunto 03	0,42	19,41	-65,31	52,42	19,34	53%
Conjunto 04	-1,01	16,81	-41,41	**34,99**	16,75	**40%**
Conjunto 05	**0,27**	16,70	-39,31	41,44	16,62	47%

Fonte: Elaborada pelo autor.

conjunto de sensores bastante aceitável.

Todos os resultados obtidos para RUL linear por partes são demonstrados na Figura 20, na qual apresenta o histograma dos erros e apresenta uma linha vertical tracejada em vermelho em zero, para servir de referência.

4.2.3 Resultado dos testes

Comparando os melhores resultados de ambos os métodos de degradação, o método RUL linear por partes apresentou um desempenho superior, reduzindo todas as medidas estatísticas e métricas de desempenho, a média dos erros por exemplo, em todos os conjuntos de sensores se aproximaram de zero, reduzindo também o desvio padrão, o valor de RMSE e o valor percentual de erros acima de zero.

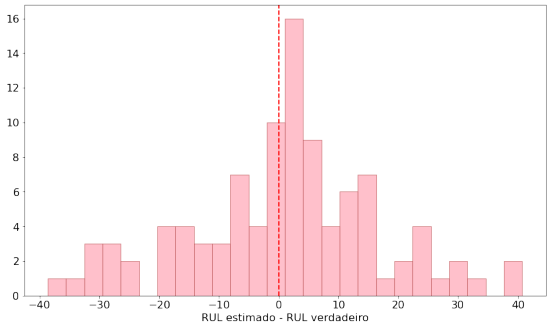
O mais importante não é apenas a questão de ter um bom desempenho nas previsões, mas também realizar a parada de manutenção do equipamento no momento certo. Por isso, as previsões iniciais são preferíveis, pois isso pode ser considerado uma avaliação positiva, no que diz respeito ao equilíbrio entre os custos de manutenção e a segurança das aeronaves, do que as previsões tardias.

Segundo Teles (2019) o objetivo principal de um plano de manutenção é minimizar o impacto de eventos não planejados em segurança, meio ambiente e rentabilidade comercial. A ferramenta de prognóstico serve para alcançar e sustentar as metas do setor de manutenção.

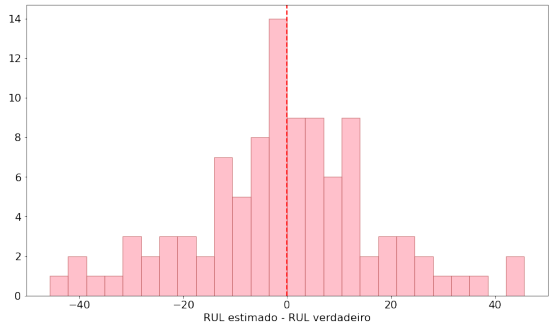
Um prognóstico bem feito e bem implementado gera diversos benefícios para o setor da manutenção, dentre eles:

- Redução dos custos de manutenção;
- Elevação dos níveis de segurança das instalações;
- Redução das atividades que não tem valor agregado;
- Melhoria na qualidade dos serviços realizados.

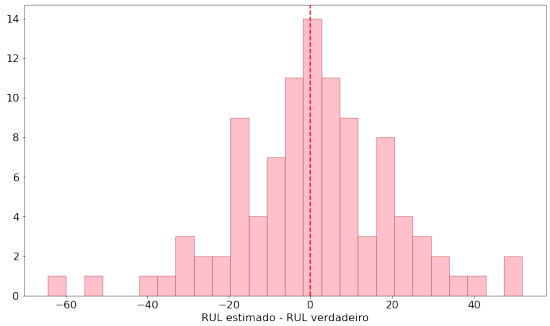
(a) Conjunto 01



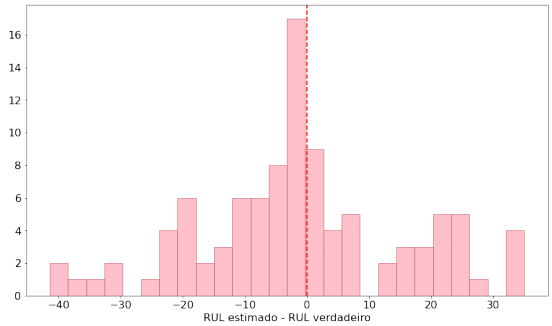
(b) Conjunto 02



(c) Conjunto 03



(d) Conjunto 04



(e) Conjunto 05

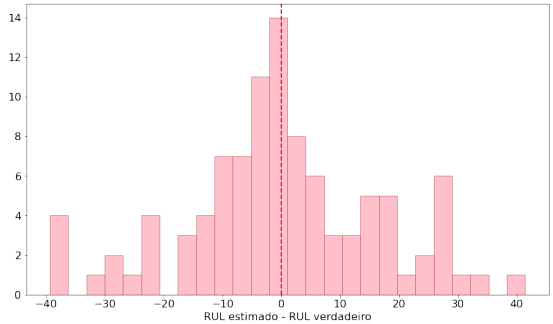


Figura 20 – Distribuição dos erros das estimativas do RUL para CNN 1D com RUL linear por partes

4.3 *Benchmarking*

Os resultados experimentais do método proposto foram comparados com os de métodos relacionados. As comparações, conforme mostrado na Tabela 6, são realizadas em termos de RMSE e a Pontuação S.

Tabela 6 – Resultados de desempenhos de outros autores.

Autor	Método	RMSE	Pontuação S
Jayasinghe et al. (2019)	Convolução temporal e LSTM	23,57	1220,00
Hsu e Jiang (2018)	RNN com duas camadas LSTM	16,74	388,68
Berghout et al. (2020b)	RT-LCI-ELM	46,03	-
Al-Dulaimi et al. (2019)	HDNN	**13,02**	245,00
Jiang e Kuo (2017)	ECNN	-	661,00
Elsheikh, Yacout e Ouali (2019)	LSTM bidirecional	63,00	376,64
Berghout et al. (2020a)	Estimativa TD baseada em DFF	13,74	**199,74**
Babu, Zhao e Li (2016)	CNN	18,45	1286,70
Zheng et al. (2018)	TW-ELM	13,78	267,31
Louen, Ding e Kandler (2013)	SVM	29,82	-
	Método proposto	16,06	424,28

Fonte: Elaborada pelo autor.

As comparações demonstram que o método proposto se equipara significativamente com outros métodos relatados na literatura apresentando Pontuação S de 424,28 e RMSE de 16,06.

4.4 Síntese do capítulo

O estudo de prognóstico de falhas foi realizado utilizando um conjunto de dados de degradação de uma turbina turbofan, o conjunto de treinamento são dados RTF, ou seja, no último ciclo o sistema apresentou uma falha funcional. Todo o processamento foi implementado no Google Colaboratory, sendo um serviço de armazenamento em nuvem de notebooks executado na nuvem, onde é possível escrever e executar códigos em Python, inicialmente mostra-se os resultados obtidos utilizando uma CNN 1D com função alvo RUL linear e variando o conjunto de sensores baseado na literatura, e depois a abordagem CNN 1D com função alvo RUL com aproximação linear por partes.

Os resultados obtidos demonstram que a abordagem com função alvo RUL linear por partes supera os resultados alcançados pela função alvo RUL linear.

5 CONCLUSÕES E TRABALHOS FUTUROS

Nesta seção são apresentados as conclusões do trabalho e as propostas para trabalhos futuros.

5.1 Conclusões

Este trabalho apresentou uma abordagem para prognóstico de falhas baseado em CNN 1D utilizando dois métodos de função alvo RUL: a linear e a aproximação linear por partes para estimar RUL. Estas abordagens criam uma função alvo de RUL similar ao método de degradação das turbinas turbofan. O método CNN 1D mais a aproximação linear por partes, obteve a melhor performance quantitativa quando comparada com a degradação linear.

A combinação do algoritmo CNN 1D com a função de aproximação linear por partes, busca simultaneamente pelos melhores padrões das séries temporais e aproxima o método de degradação a função alvo RUL. As vantagens da abordagem apresentada são que o projeto da CNN não depende da experiência do analista, já que ele permite o ajuste automático combinando diferentes parâmetros e pesos, mostrando bons resultados nas análises do subconjunto de dados "FD001"do repositório de prognóstico da NASA.

No processo experimental a comparação quantitativa através de métricas de prognóstico foi realizada com os resultados de diversas topologias, todos para o mesmo conjunto de dados, entre uma delas a CNN apresentada por Babu, Zhao e Li (2016) onde através do treinamento obteve um valor de $RMSE = 18,45$, e os melhores resultados neste trabalho para CNN 1D com RUL linear é $RMSE = 29,00$ e com a aproximação linear por partes um $RMSE = 16,06$, mostrando que com uma boa aproximação do método de degradação é possível fornecer um bom desempenho na topologia de CNN 1D.

Em geral a abordagem desenvolvida conseguiu ajustar os parâmetros de forma eficiente ao conjunto de dados, por fim, de acordo com o trabalho realizado até o momento, pode-se inferir que a estimativa de vida útil remanescente para prognóstico de falhas baseada em conjuntos de dados de simulação de degradação de turbinas turbofan se demonstra propício, pois a partir de dados e indicadores oferecidos pelos sistemas de sensoriamento, é possível identificar as pequenas irregularidades, que podem evoluir para grandes falhas, possibilitando, a redução de custos da manutenção, permitindo uma correção imediata, e assegurando com que o desempenho das máquinas e equipamentos sejam aproveitados de forma eficaz. Sendo assim, este trabalho apresenta uma contribuição para a área de prognósticos, em que estimativas antecipadas podem favorecer as atividades de inspeção e reparo de componentes.

5.2 Trabalhos futuros

Embora a premissa deste trabalho seja definido com o intuito de utilizar uma CNN 1D para resolver o problema específico de estimativa de RUL e utilizar dois modelos que representem a degradação da turbinas turbofan, existe ainda alguns problemas que necessitam ser considerados. Alguns desafios foram identificados, onde vários fatores ainda podem melhorar o resultados, as perspectivas futuras são descritas a seguir:

- Realizar um estudo de componentes principais, sendo um procedimento matemático usado para diminuir um conjunto de variáveis, determinando quais estão atuando de maneira significativa na falha observada.
- Desenvolver outros métodos de degradação temporal do equipamento, por exemplo, um modelo exponencial.
- As variáveis podem passar por um pré processamento mais significativo, por exemplo, passar por um processo de filtragem para eliminar o ruído.
- Deve-se dizer que o modelo escolhido de CNN 1D pode não ser o melhor, outros modelos podem levar em consideração outros aspectos mais significativos, podendo assim estabelecer numericamente menores métrica de desempenho.

Além disso, em um determinado projeto, as métricas de desempenho (RMSE e Pontuação S) podem não ser o objetivo principal, mas a compreensão dos recursos que explicam a falha da turbina pode ser mais importante. Esse objetivo pode ser alcançado usando algoritmos que podem fornecer uma visão sobre a importância das variáveis.

REFERÊNCIAS

ACADEMY, D. S. **Deep Learning Book**. 2021. Disponível em: <https://www.deeplearningbook.com.br/>. Acesso em: 29 abril 2021.

AL-DULAIMI, A. et al. Hybrid deep neural network model for remaining useful life estimation. In: IEEE. **ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)**. [S.l.], 2019. p. 3872–3876.

AMODEI, D. et al. Deep speech 2 : End-to-end speech recognition in english and mandarin. In: BALCAN, M. F.; WEINBERGER, K. Q. (Ed.). **Proceedings of The 33rd International Conference on Machine Learning**. New York, New York, USA: PMLR, 2016. (Proceedings of Machine Learning Research, v. 48), p. 173–182. Disponível em: <http://proceedings.mlr.press/v48/amodei16.html>.

ARAÚJO, D. C. et al. sistemas de monitoramento e diagnóstico de transformadores de subestações. **Décimo Segundo Encontro Regional Ibero-americano do CIGRÉ, Foz do Iguaçu-Pr, Brasil-20 a**, v. 24, 2007.

AVCI, O. et al. Structural damage detection in real time: implementation of 1d convolutional neural networks for shm applications. In: **Structural Health Monitoring & Damage Detection, Volume 7**. [S.l.]: Springer, 2017. p. 49–54.

BABU, G. S.; ZHAO, P.; LI, X.-L. Deep convolutional neural network based regression approach for estimation of remaining useful life. In: SPRINGER. **International conference on database systems for advanced applications**. [S.l.], 2016. p. 214–228.

BAILEY, C. et al. Current status of prognostics techniques and application to power electronics. In: IEEE. **2010 6th International Conference on Integrated Power Electronics Systems**. [S.l.], 2010. p. 1–6.

BARBOSA, L. H. A. et al. Teste para prognósticos de falhas em motores de indução: Discussão teórica e estudo de caso. **Anais da Sociedade Brasileira de Automática**, v. 2, n. 1, 2020.

BENGTSSON, M. Condition based maintenance system technology–where is development heading. **Condition Based Maintenance Systems–An Investigation of Technical Constituents and Organizational Aspects**, v. 55, 2004.

BERGHOUT, T. et al. Aircraft engines remaining useful life prediction with an improved online sequential extreme learning machine. **Applied Sciences**, Multidisciplinary Digital Publishing Institute, v. 10, n. 3, p. 1062, 2020.

BERGHOUT, T. et al. Regularized length changeable extreme learning machine with incremental learning enhancements for remaining useful life prediction of aircraft engines. In: IEEE. **020 1st International Conference on Communications, Control Systems and Signal Processing (CCSSP)**. [S.l.], 2020. p. 358–363.

BIAGETTI, T.; SCIUBBA, E. Automatic diagnostics and prognostics of energy conversion processes via knowledge-based systems. **Energy**, Elsevier, v. 29, n. 12-15, p. 2553–2572, 2004.

BROWNLEE, J. **A gentle introduction to pooling layers for convolutional neural networks.** 2019. Disponível em: <https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/>. Acesso em: 30 abril 2021.

COLLOBERT, R.; WESTON, J. A unified architecture for natural language processing: Deep neural networks with multitask learning. In: **Proceedings of the 25th international conference on Machine learning**. [S.l.: s.n.], 2008. p. 160–167.

CONNOR, J. T.; MARTIN, R. D.; ATLAS, L. E. Recurrent neural networks and robust time series prediction. **IEEE transactions on neural networks**, IEEE, v. 5, n. 2, p. 240–254, 1994.

DJEZIRI, M. et al. Fault diagnosis and prognosis based on physical knowledge and reliability data: Application to mos field-effect transistor. **Microelectronics Reliability**, Elsevier, v. 110, p. 113682, 2020.

ELSHEIKH, A.; YACOUT, S.; OUALI, M.-S. Bidirectional handshaking lstm for remaining useful life prediction. **Neurocomputing**, Elsevier, v. 323, p. 148–156, 2019.

FUKUSHIMA, K.; MIYAKE, S. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. **Pattern recognition**, Elsevier, v. 15, n. 6, p. 455–469, 1982.

GAD, A. **A Comprehensive Guide to the Backpropagation Algorithm in Neural Networks**. 2021. Disponível em: <https://neptune.ai/blog/backpropagation-algorithm-in-neural-networks-guide>. Acesso em: 19 maio 2021.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. [S.l.]: MIT Press, 2016. <http://www.deeplearningbook.org>.

GOURIVEAU, R.; ZERHOUNI, N. Connexionist-systems-based long term prediction approaches for prognostics. **IEEE Transactions on Reliability**, IEEE, v. 61, n. 4, p. 909–920, 2012.

GRANAT, M. **How to use convolutional neural networks for time series classification.** 2019. Disponível em: <https://towardsdatascience.com/how-to-use-convolutional-neural-networks-for-time-series-classification-56b1b0a07a57>. Acesso em: 02 maio 2021.

HOU, M.; PI, D.; LI, B. Similarity-based deep learning approach for remaining useful life prediction. **Measurement**, Elsevier, v. 159, p. 107788, 2020.

HSU, C.-S.; JIANG, J.-R. Remaining useful life estimation using long short-term memory deep learning. In: IEEE. **2018 IEEE International Conference on Applied System Invention (ICASI)**. [S.l.], 2018. p. 58–61.

INCE, T. et al. Real-time motor fault detection by 1-d convolutional neural networks. **IEEE Transactions on Industrial Electronics**, IEEE, v. 63, n. 11, p. 7067–7075, 2016.

ISHIBASHI, R.; JÚNIOR, C. L. N. Gfrbs-phm: A genetic fuzzy rule-based system for phm with improved interpretability. In: IEEE. **2013 IEEE Conference on Prognostics and Health Management (PHM)**. [S.l.], 2013. p. 1–7.

JAYASINGHE, L. et al. Temporal convolutional memory networks for remaining useful life estimation of industrial machinery. In: IEEE. **2019 IEEE International Conference on Industrial Technology (ICIT)**. [S.l.], 2019. p. 915–920.

JIANG, J.-R.; KUO, C.-K. Enhancing convolutional neural network deep learning for remaining useful life estimation in smart factory applications. In: IEEE. **2017 International Conference on Information, Communication and Engineering (ICICE)**. [S.l.], 2017. p. 120–123.

JORDAN, J. **Convolutional neural networks.** 2017. Disponível em: <https://www.jeremyjordan.me/convolutional-neural-networks/>. Acesso em: 02 maio 2021.

KIRANYAZ, S. et al. 1d convolutional neural networks and applications: A survey. **Mechanical Systems and Signal Processing**, Elsevier, v. 151, p. 107398, 2021.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. **Advances in neural information processing systems**, v. 25, p. 1097–1105, 2012.

KıZRAK, A. **Comparison of Activation Functions for Deep Neural Networks**. 2019. Disponível em: <https://towardsdatascience.com/comparison-of-activation-functions-for-deep-neural-networks-706ac4284c8a>. Acesso em: 02 maio 2021.

LECUN, Y. et al. Handwritten digit recognition with a back-propagation network. In: **Advances in neural information processing systems**. [S.l.: s.n.], 1990. p. 396–404.

LEE, J. et al. Intelligent prognostics tools and e-maintenance. **Computers in industry**, Elsevier, v. 57, n. 6, p. 476–489, 2006.

LEVITT, J. **Complete guide to preventive and predictive maintenance**. [S.l.]: Industrial Press Inc., 2003.

LI, L. **Introduction to Multilayer Neural Networks with TensorFlow's Keras API**. 2019. Disponível em: <https://towardsdatascience.com/introduction-to-multilayer-neural-networks-with-tensorflows-keras-api-abf4f813959>. Acesso em: 29 abril 2021.

LI, X.; DING, Q.; SUN, J.-Q. Remaining useful life estimation in prognostics using deep convolution neural networks. **Reliability Engineering & System Safety**, Elsevier, v. 172, p. 1–11, 2018.

LIAO, L.; KÖTTIG, F. Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction. **IEEE Transactions on Reliability**, IEEE, v. 63, n. 1, p. 191–207, 2014.

LIU, K.; GEBRAEEL, N. Z.; SHI, J. A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis. **IEEE Transactions on Automation Science and Engineering**, IEEE, v. 10, n. 3, p. 652–664, 2013.

LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. **CoRR**, abs/1411.4038, 2014.

LOUEN, C.; DING, S.; KANDLER, C. A new framework for remaining useful life estimation using support vector machine classifier. In: IEEE. **2013 Conference on Control and Fault-Tolerant Systems (SysTol)**. [S.l.], 2013. p. 228–233.

LUO, J. et al. Model-based prognostic techniques [maintenance applications]. In: IEEE. **Proceedings AUTOTESTCON 2003. IEEE Systems Readiness Technology Conference.** [S.l.], 2003. p. 330–340.

MATHWORKS. **What Is Deep Learning?** 2021. Disponível em: <https://www.mathworks.com/discovery/deep-learning.html>. Acesso em: 29 abril 2021.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, Springer, v. 5, n. 4, p. 115–133, 1943.

MEDJAHER, K.; ZERHOUNI, N.; BAKLOUTI, J. Data-driven prognostics based on health indicator construction: Application to pronostia's data. In: IEEE. **2013 European Control Conference (ECC)**. [S.l.], 2013. p. 1451–1456.

NASA. **Prognostics Center of Excellence**. 2021. Disponível em: <https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/>. Acesso em: 24 fev 2021.

OPPERMANN, A. **What is Deep Learning and How does it work?** 2019. Disponível em: <https://towardsdatascience.com/what-is-deep-learning-and-how-does-it-work-2ce44bb692ac>. Acesso em: 29 abril 2021.

PETRILLO, A. et al. Model-based vehicular prognostics framework using big data architecture. **Computers in Industry**, v. 115, p. 103177, 2020. ISSN 0166-3615. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0166361519304439>.

RAMASSO, E.; ROMBAUT, M.; ZERHOUNI, N. Joint prediction of continuous and discrete states in time-series based on belief functions. **IEEE transactions on cybernetics**, IEEE, v. 43, n. 1, p. 37–50, 2012.

RAMASSO, E.; SAXENA, A. Review and analysis of algorithmic approaches developed for prognostics on cmapss dataset. In: **Annual Conference of the Prognostics and Health Management Society 2014.** [S.l.: s.n.], 2014.

RANZATO, M. et al. Efficient learning of sparse representations with an energy-based model. **Advances in neural information processing systems**, MIT; 1998, v. 19, p. 1137, 2007.

REYES, K. **What is Deep Learning and How Does It Works**. 2021. Disponível em: <https://www.simplilearn.com/tutorials/deep-learning-tutorial/what-is-deep-learning>. Acesso em: 29 abril 2021.

RIZWAN, M. **Implementation of Convolutional Neural Network Using Keras.** 2018. Disponível em: <https://towardsdatascience.com/implementation-of-convolutional-neural-network-using-keras-9ace67916c07>. Acesso em: 02 maio 2021.

SAHOO, B. **Data-Driven Remaining Useful Life (RUL) Prediction**. 2018. Disponível em: <https://biswajitsahoo1111.github.io/rul_codes_open/>.

SANTOS, A. et al. Uma abordagem de classificação de imagens dermatoscópicas utilizando aprendizado profundo com redes neurais convolucionais. In: SBC. **Anais do XVII Workshop de Informática Médica**. [S.l.], 2017.

SARKAR, S.; JIN, X.; RAY, A. Data-driven fault detection in aircraft engines with noisy sensor measurements. **Journal of Engineering for Gas Turbines and Power**, American Society of Mechanical Engineers Digital Collection, v. 133, n. 8, 2011.

SAXENA, A. et al. Damage propagation modeling for aircraft engine run-to-failure simulation. In: IEEE. **2008 international conference on prognostics and health management**. [S.l.], 2008. p. 1–9.

SIKORSKA, J.; HODKIEWICZ, M.; MA, L. Prognostic modelling options for remaining useful life estimation by industry. **Mechanical systems and signal processing**, Elsevier, v. 25, n. 5, p. 1803–1836, 2011.

SIMEÓN, E. J. A. Prognóstico de falhas baseado em redes neurais com estados de eco. 2015.

SÖDERHOLM, P. **Maintenance and continuous improvement of complex systems: linking stakeholder requirements to the use of built-in test systems**. Tese (Doutorado) — Luleå tekniska universitet, 2005.

TEIXEIRA, L. M. et al. Prognóstico e detecção de falhas em plantas industriais. 2014.

TELES, J. **Planejamento e Controle da Manutenção**. [S.l.]: Engeteles, 2019.

THURSTON, M.; LEBOLD, M. **Standards developments for condition-based maintenance systems**. [S.l.], 2001.

TODINOV, M. Reliability analysis based on the losses from failures. **Risk Analysis: An International Journal**, Wiley Online Library, v. 26, n. 2, p. 311–335, 2006.

URBAN, G. et al. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. **Gastroenterology**, v. 155, n. 4, p. 1069–1078.e8, 2018. ISSN 0016-5085. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0016508518346596>.

VACHTSEVANOS, G. et al. **Intelligent fault diagnosis and prognosis for engineering systems: vachtsevanos/intelligent fault diagnosis**. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2006. ISBN 9780470117842 9780471729990. Disponível em: <http://doi.wiley.com/10.1002/9780470117842>.

WANG, H. et al. Application of deep-learning algorithms to mstar data. In: IEEE. **2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)**. [S.l.], 2015. p. 3743–3745.

WANG, T. et al. A similarity-based prognostics approach for remaining useful life estimation of engineered systems. In: IEEE. **2008 international conference on prognostics and health management**. [S.l.], 2008. p. 1–6.

ZHENG, C. et al. A data-driven approach for remaining useful life prediction of aircraft engines. In: IEEE. **2018 21st International Conference on Intelligent Transportation Systems (ITSC)**. [S.l.], 2018. p. 184–189.

ZIHLMANN, M.; PEREKRESTENKO, D.; TSCHANNEN, M. Convolutional recurrent neural networks for electrocardiogram classification. In: IEEE. **2017 Computing in Cardiology (CinC)**. [S.l.], 2017. p. 1–4.