



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CAMPUS AGRESTE
NÚCLEO DE TECNOLOGIA
CURSO DE ENGENHARIA DE PRODUÇÃO

MARCOS VINICIUS LIRA SANTANA

**PREVISÃO DO TEMPO DE VIDA RESTANTE PARA ROLAMENTOS UTILIZANDO
UM MODELO DE REDE NEURAL CONVOLUCIONAL**

Caruaru

2024

MARCOS VINICIUS LIRA SANTANA

**PREVISÃO DO TEMPO DE VIDA RESTANTE PARA ROLAMENTOS UTILIZANDO
UM MODELO DE REDE NEURAL CONVOLUCIONAL**

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Engenharia de Produção do Campus Agreste da Universidade Federal de Pernambuco – UFPE, na modalidade de monografia, como requisito parcial para a obtenção do grau de bacharel/licenciado em Engenharia de Produção.

Área de concentração: Inteligência Artificial.

Orientador (a): Caio Bezerra Souto Maior

Caruaru

2024

MARCOS VINICIUS LIRA SANTANA

**PREVISÃO DO TEMPO DE VIDA RESTANTE PARA ROLAMENTOS UTILIZANDO
UM MODELO DE REDE NEURAL CONVOLUCIONAL**

Trabalho de Conclusão de Curso apresentado à
Coordenação do Curso de Engenharia de Produção
do Campus Agreste da Universidade Federal de
Pernambuco – UFPE, na modalidade de
monografia, como requisito parcial para a obtenção
do grau de bacharel/licenciado em Engenharia de
Produção.

Aprovada em: 12/03/2024

BANCA EXAMINADORA

Pro. Dr. Caio Bezerra Souto Maior (Orientadora)
Universidade Federal de Pernambuco

Prof. Dr. Lucimario Gois De Oliveira Silva (Examinador Interno)
Universidade Federal de Pernambuco

Prof^ª. Dr^ª. July Bias Macedo (Examinadora Interna)
Universidade Federal de Pernambuco

Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE

Santana, Marcos Vinicius Lira.

Previsão do tempo de vida restante para rolamentos utilizando um modelo de rede neural convolucional / Marcos Vinicius Lira Santana. - Caruaru, 2024.
38p : il., tab.

Orientador(a): Caio Bezerra Souto Maior

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro Acadêmico do Agreste, Engenharia de Produção, 2024.

Inclui referências, apêndices, anexos.

1. manutenção. 2. rolamentos. 3. redes neurais convolucionais. I. Maior, Caio Bezerra Souto. (Orientação). II. Título.

620 CDD (22.ed.)

RESUMO

Em ambientes industriais, a eficiência operacional e a manutenção proativa são pilares essenciais para garantir a produtividade e a confiabilidade dos equipamentos. No âmbito desse contexto, a previsão do Remaining Useful Life (RUL), ou vida útil restante, de componentes como rolamentos industriais desempenha um papel fundamental. Ao antecipar o momento ideal para substituição ou reparo desses componentes, as empresas podem evitar falhas inesperadas, reduzir custos de manutenção e maximizar a disponibilidade operacional dos equipamentos. Este trabalho de conclusão de curso concentra-se em desenvolver e avaliar um modelo de previsão do RUL para rolamentos industriais, utilizando uma abordagem baseada em Deep Learning, mais especificamente uma Rede Neural Convolucional (CNN). A escolha por uma abordagem de Deep Learning se justifica pela sua capacidade de extrair padrões complexos e não-lineares a partir de grandes volumes de dados, o que pode levar a previsões mais precisas e confiáveis. A metodologia adotada neste estudo abrangeu desde a coleta de dados reais de rolamentos até o processamento, seleção de variáveis relevantes, implementação e treinamento da CNN. O objetivo principal é prever o RUL para possibilitar a otimização dos processos de manutenção e maximizar a disponibilidade dos equipamentos. Isto é demonstrado através da eficácia e métricas de avaliação do desempenho dos modelos treinados. Os resultados obtidos demonstram a eficácia da abordagem proposta. O modelo de previsão do RUL para os rolamentos industriais alcançou um Mean Absolute Percentage Error (MAPE) de 53%, indicando uma precisão notável na estimativa do tempo de vida útil restante. Em comparação com um estudo anterior (LI, 2019), que previa um MAPE de 62,3% para o mesmo tipo de rolamento e condições. Espera-se que os resultados experimentais deste estudo evidenciem a eficácia da abordagem proposta, comparando-os com referências da literatura. A validação do modelo em cenários reais de operação de rolamentos industriais será fundamental para demonstrar sua utilidade e potencial aplicabilidade em contextos industriais.

Palavras-chave: manutenção; rolamentos; redes neurais convolucionais.

ABSTRACT

In industrial environments, operational efficiency and proactive maintenance are essential pillars to ensure equipment productivity and reliability. Within this context, predicting the Remaining Useful Life (RUL) of components such as industrial bearings plays a fundamental role. By anticipating the ideal time for replacement or repair of these components, companies can avoid unexpected failures, reduce maintenance costs, and maximize equipment operational availability. This thesis focuses on developing and evaluating an RUL prediction model for industrial bearings using a Deep Learning approach, specifically a Convolutional Neural Network (CNN). The choice of Deep Learning is justified by its ability to extract complex and non-linear patterns from large volumes of data, potentially leading to more precise and reliable predictions. The methodology adopted in this study ranged from collecting real bearing data to processing, selecting relevant variables, implementing, and training the CNN. The main objective is to predict RUL to optimize maintenance processes and maximize equipment availability. This is demonstrated through the effectiveness and performance evaluation metrics of the trained models. The results obtained show the effectiveness of the proposed approach. The RUL prediction model for industrial bearings achieved a Mean Absolute Percentage Error (MAPE) of 53%, indicating notable accuracy in estimating remaining useful life. In comparison with a previous study (LI, 2019) predicting a MAPE of 62.3% for the same type of bearing and conditions, our approach shows improvement. It is expected that the experimental results of this study will highlight the effectiveness of the proposed approach by comparing them with literature references. Validating the model in real operational scenarios of industrial bearings will be crucial to demonstrate its utility and potential applicability in industrial contexts.

Keywords: maintenance; bearings; convolutional neural networks.

SUMÁRIO

1	INTRODUÇÃO	7
1.1.	OBJETIVO	8
1.1.1.	Geral	8
1.1.2.	Específico.....	8
2	REFERENCIAL TEÓRICO	10
2.1.	CONFIABILIDADE E MANUTENÇÃO INDUSTRIAL	10
2.1.1.	RUL.....	12
2.1.2.	Rolamentos.....	13
2.2	ESPECTROGRAMA	14
2.3.	IA.....	16
2.3.1.	Machine Learning	16
2.3.2.	Deep Learning.....	18
2.3.3.	CNN	19
3	CONJUNTO DE DADOS	23
4	METODOLOGIA	26
4.1.	COLETA DE DADOS	26
4.2.	PRÉ-PROCESSAMENTO DE DADOS.....	26
4.3	SELEÇÃO E PREPARAÇÃO DE VARIÁVEIS	27
4.4.	MODELAGEM	27
4.5.	AVALIAÇÃO DE RESULTADOS	29
5	RESULTADOS	30
6	CONCLUSÃO	34
	REFERÊNCIAS	35

1 INTRODUÇÃO

A maioria dos custos, cerca de 70-80% de um ativo de instalação é realizada durante a fase de operação e manutenção da vida útil da instalação ciclo. Um estudo do Instituto Nacional de Padrões e Tecnologia (NIST) de 2018 resumiu que US\$ 50 bilhões são gastos anualmente em operações e manutenção no território continental dos Estados Unidos (SCAIFE et al., 2024).

As máquinas são os principais componentes de qualquer unidade industrial ou fábrica. Assim, o lucro de qualquer indústria está altamente associado ao tempo de execução disponível das máquinas. O colapso repentino dos componentes da máquina pode levar a perdas substanciais de produção. O monitoramento adequado das condições desses componentes é essencial para garantir o funcionamento ininterrupto das indústrias (SAHA et al., 2022).

Ambientes industriais são constituídos por diferentes tipos de equipamentos que fazem parte de sistemas de produção complexos, dentre os quais um dos mais importantes, e mais sujeitos a falhas, são as máquinas rotativas. Dessa forma a utilização de técnicas de manutenção eficientes são importantes para que esses equipamentos estejam sempre em condições de produção (SOARES et al., 2020).

Apesar da sua indiscutível relevância, os processos de manutenção frequentemente se deparam com desafios ligados à previsibilidade, custo e impactos na produção. Os métodos tradicionais, em sua maioria baseados em manutenções periódicas ou reativas, muitas vezes são ineficientes, levando à subutilização de recursos e à detecção insuficiente de falhas incipientes. Dessa forma, redes neurais artificiais, um modelo de Inteligência artificial (IA) têm sido utilizadas em vários problemas de monitoração de condição, diagnóstico e prognóstico de falhas em máquinas (LUCIFREDI et al., 2000).

ZHANG et al., 2017 afirma que modelos de diagnóstico de falhas baseados em métodos orientados a dados são vantajosos, pois podem fornecer diagnósticos precisos, mesmo sem o conhecimento do modelo do processo. Com a crescente demanda de monitoramento e a grande quantidade de informações de dados de processos industriais, a utilização de IA ganha cada vez mais atenção pois, por meio destas técnicas, é possível a extração de informações úteis de dados históricos (CHEN et al., 2019).

Avanços recentes em IA na indústria mostraram o potencial dessa tecnologia para auxiliá-las a enfrentar os desafios associados a essa transformação digital de Sistemas Ciber físicos (PERES et al., 2020). Clique ou toque aqui para inserir o texto. Combinando análises complexas com a

capacidade de adaptação a diferentes contextos, a IA supera as abordagens tradicionais, viabilizando um paradigma proativo de manutenção que, por sua vez, impacta positivamente os custos operacionais e a disponibilidade dos equipamentos.

Nesse contexto, o tempo de vida útil restante (RUL) é uma métrica importante. O 'Businessdictionary.com' define a vida útil como 'o período durante o qual se espera que um ativo ou propriedade seja utilizável para a finalidade para a qual foi adquirido'. No entanto, em engenharia de confiabilidade, ela pode ser formalmente definida como 'o período de tempo esperado durante o qual um ativo depreciado será produtivo' (SI et al., 2011).

O RUL fornece suporte teórico e base de tomada de decisão para a gestão de sistemas de saúde. Uma vez prevista a falha, as decisões de manutenção serão implementadas a tempo de reduzir o risco de perdas econômicas, garantindo ao mesmo tempo a segurança, a eficiência da manutenção e a confiabilidade do sistema (CAI et al., 2021).

Utilizar na estimação do momento de falha de um equipamento traz vantagens competitivas ao definir o momento apropriado para manutenção. A crescente escassez e o consequente aumento dos preços dos recursos bem como a mudança social no contexto dos aspectos ecológicos exigem um foco mais intenso na sustentabilidade em contextos industriais (LASI et al., 2014).

Nesse contexto, este trabalho busca explorar modelos de IA dentro da engenharia, focando na previsão do tempo de vida remanescente de rolamentos industriais. Especificamente, técnicas de *Machine Learning (ML)* e *Deep Learning (DP)* serão aplicadas em um conjunto de dados reais disponibilizado no IEEE PHM 2012 Prognostic challenge.

1.1 OBJETIVO

1.1.1 Geral

Apresentar um modelo de previsão do RUL para rolamentos utilizando uma abordagem de CNN com intuito de fornecer previsões mais precisas que poderá dar suporte ao planejamento da manutenção.

1.1.2 Específico

- Fazer revisão da literatura sobre os seguintes tópicos:
 - Manutenção;
 - Confiabilidade;
 - RUL;

- CNN;
- SPECTOGRAMA;
- Coletar dados de vibração do IEEE PHM 2012 Prognostic challenge;
- Realizar um pré-processamento para adequação dos dados utilizados;
- Seleção de variáveis informativas para dados de vibração;
- Modelagem prognóstica via uma abordagem CNN;
- Comparação de resultados.

2 REFERENCIAL TEÓRICO

2.1 CONFIABILIDADE E MANUTENÇÃO INDUSTRIAL

A manutenção tem um papel essencial na garantia da qualidade e da produtividade empresarial. A manutenção deve ser encarada como uma função estratégica na obtenção dos resultados da organização e deve estar direcionada ao suporte do gerenciamento e à solução de problemas apresentados na produção, lançando a empresa em patamares competitivos de qualidade e produtividade (KARDEC & NASCIF, 2001).

O custo da indisponibilidade concentra-se naqueles decorrentes da perda de produção, da não-qualidade dos produtos, da recomposição da produção e das penalidades comerciais, com possíveis consequências sobre a imagem da empresa (MIRSHAWKA; 1993).

A manutenção é uma combinação de ações técnicas, incluindo as de verificação, destinadas a manter ou reparar um bem de um equipamento, para que possa desempenhar a sua função. Sendo assim, as principais práticas básicas de manutenção são: Manutenção corretiva, manutenção preventiva, manutenção preditiva e manutenção detectiva (ARMANDO DE ALMEIDA MURÇA et al., 2012).

- **Manutenção corretiva:** é a manutenção que acontece após a identificação de um problema. O problema dessa política não está em fazer intervenções corretivas, mas em que sua aplicação isolada requer enormes estoques de peças para suportar as sucessivas quebras, tornando o trabalho imprevisível e, portanto, sem um plano capaz de equacionar os custos.
- **Manutenção preventiva:** baseia-se em intervenções periódicas geralmente programadas segundo a frequências definida pelos fabricantes dos equipamentos. Essa política, em muitos casos, leva a desperdícios, pois não considera a condição real do equipamento.
- **Manutenção preditiva:** caracteriza-se pela medição e análise de variáveis da máquina que possam prognosticar uma eventual falha. Com isso, a equipe de manutenção pode se programar para a intervenção e aquisição de peças (custo da manutenção), reduzindo gastos com estoque e evitando paradas desnecessárias da linha de produção (custo da indisponibilidade).

A manutenção preditiva é indicada para equipamento de altos custos onde o monitoramento do sistema permite o entendimento de seu funcionamento e criação de planos de ação para evitar problemas futuros. Análise de dados como vibração e temperatura podem informar sobre o estado

de degradação de um ativo e, a partir daí, podem tomar decisões sobre como lidar com o processo produtivo.

Geralmente, condições anormais de funcionamento de rolamentos como o desgaste, se manifestam na forma de vibração, o monitoramento dessa variável é uma das principais ferramentas para acompanhamento das condições de funcionamento das máquinas (SOARES et al., 2020).

Neste contexto, confiabilidade pode ser entendida como a probabilidade de um sistema realizar tarefas para ele foi projetado. Sob certas condições, durante um período de tempo predefinido (HOYLAND & RAUSAND, 2009). Ou seja, quando se quer especificar a confiabilidade de uma máquina, deve-se informar que “a probabilidade dessa máquina operar de acordo com suas especificações nesse projeto é, por exemplo, de 99% nas próximas 5000 horas’ e não somente ‘A confiabilidade dessa máquina é de 99%’. O estudo dessa probabilidade permite antever eventos de falhas em um sistema e tomar decisões que mitiguem ou evitem esses eventos, considerados indesejados.

Ainda, os termos “disponibilidade” e “manutenibilidade” (ou manutenibilidades) podem ser definidos como abaixo (NBR 5462:1994).

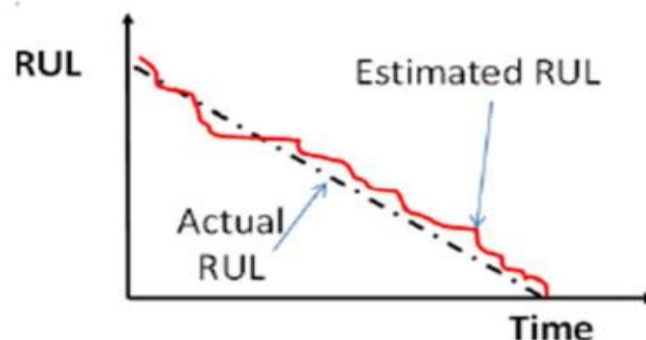
- **Disponibilidade:** É a capacidade de um item estar em condições de executar uma certa função em um dado instante ou durante um intervalo de tempo determinado, levando-se em conta os aspectos combinados de sua confiabilidade, manutenibilidade e suporte de manutenção, supondo que os recursos externos requeridos estejam assegurados.
- **Manutenibilidade:** é a capacidade de um item ser mantido ou recolocado em condições de executar suas funções requeridas, sob condições de uso especificadas, quando a manutenção é executada sob condições determinadas e mediante procedimentos e meios prescritos.” Em outras palavras, a manutenibilidade é a rapidez com que procedimentos de reparo são executados para recolocar um ativo em funcionamento após a apresentação de um problema.

Em outras palavras, disponibilidade é um indicador reativo, retrata o que aconteceu no passado. Manutenibilidade é a “facilidade” que uma equipe encontra para executar a manutenção de um determinado equipamento. A confiabilidade está ligado ao futuro, é uma projeção probabilística que aponta as chances de o equipamento funcionar perfeitamente em um determinado espaço de tempo.

2.1.1 RUL

Como mencionado anteriormente, o conceito de RUL pertence às áreas de Engenharia e IA e se refere ao período de tempo estimado que um componente, máquina ou sistema possui até que alcance um estado de degradação que não permita mais seu funcionamento eficiente. A Figura 1 demonstra graficamente a comparação entre um RUL real e sua respectiva previsão.

Figura 1: Comportamento esperado do RUL estimado



Fonte: (SUTRISNO et al., 2012)

O RUL é uma informação utilizada na área de manutenção preditiva e prognóstico para estimar a quantidade de tempo restante antes de uma falha ou degradação significativa de um componente, equipamento ou sistema. O RUL é uma medida importante para prever a falha ou a necessidade de manutenção com antecedência, permitindo uma abordagem mais proativa e eficiente na gestão dos ativos (SAXENA et al., 2010).

O RUL desempenha um papel crucial na manutenção preditiva e na otimização dos processos de produção, especialmente em setores com equipamentos complexos e caros, como aeronáutica, energia, manufatura e transporte. Ao estimar o tempo restante de vida útil dos ativos, as empresas podem programar suas atividades de manutenção de forma mais eficiente, evitando paradas inesperadas e reduzindo custos associados a falhas e reparos de emergência. Além disso, o RUL permite otimizar o uso dos recursos, estender a vida útil dos ativos e planejar estrategicamente as substituições quando necessário (SAXENA et al., 2010).

O RUL é uma métrica fundamental para a manutenção preditiva, pois fornece informações valiosas sobre o estado atual dos ativos e a expectativa de sua vida útil restante. Com base no RUL, é possível tomar decisões informadas sobre quando realizar a manutenção preventiva, programar intervalos de manutenção, gerenciar estoques de peças de reposição e até mesmo planejar

investimentos em novos equipamentos. Além disso, o RUL pode ser usado como um indicador-chave de desempenho (KPI) para avaliar a eficácia de estratégias de manutenção e comparar o desempenho de diferentes ativos ou sistemas (SI et al., 2011).

2.1.2 Rolamentos

Motores são um dos componentes principais para o funcionamento das máquinas industriais. As estatísticas mostram que os motores elétricos são usados como motores principais em mais de 90% dos acionamentos mecânicos. (ZU, 2019) A falha nesses motores geralmente cai na categoria de falhas relacionadas ao rolamento, falhas associadas ao enrolamento do estator e falhas relacionadas ao rotor. Dentre essas falhas, a maior parte delas, quase 40%, ocorre devido a problemas relacionados aos rolamentos (MOHANTY, 2014).

Os rolamentos estão entre os componentes mais importantes e frequentemente encontrados na grande maioria das máquinas rotativas, sendo sua capacidade de carga e confiabilidade proeminentes para o desempenho geral da máquina. Portanto, naturalmente, a identificação de falhas em rolamentos de elementos rolantes tem sido objeto de extensa pesquisa (TANDON; INTERNATIONAL; 1999). A Figura 2 ilustra rolamentos em condições normais e submetidos a desgaste, apresentando sinais visíveis de deterioração ao longo do tempo.

Figura 2: Rolamentos normais e degradados



Fonte: Adaptado de NECTOUX et al., 2012

Rolamentos são componentes básicos essenciais para uma máquina rotativa e isto motiva o estudo da detecção antecipada de problemas para que não haja impactos nos processos produtivos onde eles estão presentes (SOARES et al., 2020).

Tendo em mente que gerar dados é a base da Indústria 4.0, sensores industriais possuem cada vez mais importância no setor. Cada tipo de sensor é capaz de detectar determinada grandeza física. Em seguida, ele transforma essa grandeza em sinais elétricos, sonoros em dados (DJP AUTOMAÇÃO, 2020). Nos últimos anos, os sensores tornaram-se mais baratos e prontamente disponíveis para indústrias (SAN MARTIN et al., 2019).

Métodos diferentes são utilizados para a aquisição de informações e podem ser amplamente classificados dependendo do tipo de medições: vibração e acústica, temperatura e análise de resíduos de desgaste (TANDON & INTERNATIONAL, 1999). Dentre esses, as medições de vibração são comumente utilizadas no monitoramento de condições e diagnóstico de máquinas rotativas, principalmente devido aos sinais fáceis de mensurar e à análise plausível (SOUTO MAIOR, 2017).

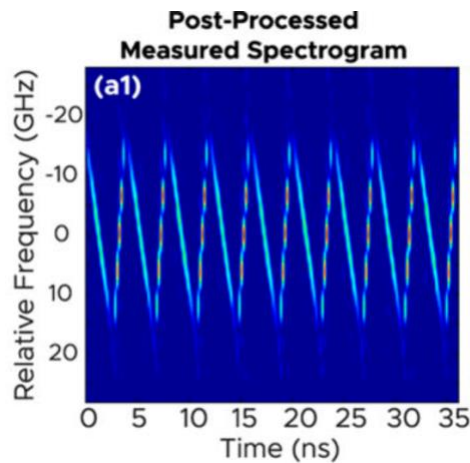
A medição de vibração do rolamento pode ser feita usando sensores de aceleração que são colocados na carcaça do rolamento. Quando ocorrem falhas no rolamento, o sinal de vibração será diferente do sinal em estado normal. Falhas localizadas em rolamentos produzem uma série de respostas de impulsos de banda larga no sinal de aceleração, à medida que os componentes do rolamento repetidamente impactam a falha (SOUTO MAIOR, 2017).

2.2 ESPECTROGRAMA

Uma das representações mais intuitivas de uma forma de onda é alcançada por meio da análise tempo-frequência, que descreve como os componentes de frequência de uma onda evoluem ao longo do tempo. Representações tempo-frequência, como o espectrograma, são bem conhecidas por permitir a caracterização completa de um sinal em termos de amplitude e fase (CROCKETT et al., 2023).

Na Figura 3 podemos visualizar o exemplo de um espectrograma. Esses espectros de frequência são então plotados em um gráfico bidimensional, onde o eixo horizontal representa o tempo e o eixo vertical representa a frequência. A intensidade ou amplitude de cada frequência é geralmente representada por uma escala de cores que vão desde tons de azul para representar baixas intensidades, passando por tons de verde e amarelo para intensidades médias, até tons de vermelho e laranja para intensidades mais altas.

Figura 3: Exemplo de um espectrograma



Fonte: CROCKETT et al., 2023

O espectrograma é uma ferramenta fundamental na análise de sinais em diversas áreas. Ele revela informações cruciais contidas em sinais complexos, permitindo a visualização e compreensão de componentes de frequência que variam com o tempo.

Espectrograma é definido como a magnitude ao quadrado da transformada de Fourier de curto tempo (STFT), que é gerada ao calcular a transformada de Fourier de seções consecutivas temporalmente truncadas (ou janeladas) do sinal de entrada, exibindo assim o seu conteúdo de frequência variável ao longo do tempo (LEON COHEN, 1995).

Espectrograma é uma representação visual do espectro de frequências do som. No seu formato mais comum, ele é representado por um gráfico em que o eixo horizontal representa o tempo e o eixo vertical a frequência (COSTA, 2013).

O espectrograma em sua essência é uma matriz. Essa matriz pode ser representada em forma de imagem, onde a entrada i, j na matriz corresponde à intensidade do pixel i, j na imagem (JANA et al., 2020).

Embora o espectrograma seja uma ferramenta poderosa, ele não está isento de desafios. A escolha de parâmetros, como tamanho da janela e taxa de sobreposição, pode afetar a resolução temporal e espectral do espectrograma. Além disso, a interpretação requer conhecimento especializado para distinguir informações relevantes de artefatos.

2.3 IA

2.3.1 Machine Learning

Na última década, houve um notável aumento nas técnicas baseadas em ML, impactando muitas áreas da indústria, incluindo direção autônoma, cuidados de saúde, finanças e manufatura. O objetivo geral do ML é reconhecer padrões em dados, que informam a maneira como problemas não vistos são tratados. Por exemplo, em um sistema altamente complexo, como um carro autônomo, grandes quantidades de dados provenientes de sensores precisam ser convertidas em decisões sobre como controlar o carro por um computador que "aprendeu" a reconhecer o padrão de "perigo" (CARLEO et al., 2019).

Normalmente, as metodologias de ML envolvem um processo de aprendizado com o objetivo de aprender com "experiência" (dados de treinamento) para realizar uma tarefa. Os dados em ML consistem em um conjunto de exemplos. Geralmente, um exemplo individual é descrito por um conjunto de atributos, também conhecidos como características ou variáveis. Uma característica pode ser nominal (enumeração), binária (ou seja, 0 ou 1), ordinal (por exemplo, A+ ou B-), ou numérica (número inteiro, número real, etc.) (LIAKOS et al., 2018).

O desempenho do modelo de ML em uma tarefa específica é medido por uma métrica de desempenho que melhora com a experiência ao longo do tempo. Para calcular o desempenho de modelos e algoritmos de ML, são usados vários modelos estatísticos e matemáticos. Após o término do processo de aprendizado, o modelo treinado pode ser usado para classificar, prever ou agrupar novos exemplos (dados de teste) usando a experiência obtida durante o processo de treinamento (LIAKOS et al., 2018). Tradicionalmente, os modelos de ML são divididos em 3 classes (RASCHKA, 2015):

- **Aprendizado supervisionado:** O aprendizado supervisionado é a tarefa de aprendizado de máquina que consiste em aprender uma função que mapeia uma entrada para uma saída com base em pares de entrada e saída de exemplo. Ele infere uma função a partir de dados de treinamento rotulados, que consistem em um conjunto de exemplos de treinamento. Os algoritmos de aprendizado de máquina supervisionado são aqueles que necessitam de assistência externa (MAHESH; 2020).

Em problemas de regressão, são fornecidas amostras de um conjunto de variáveis independentes (preditoras) x_1, x_2, \dots, x_n e o valor da respectiva variável dependente (saída) y . Nosso objetivo é obter um modelo que de alguma forma capture o mapeamento $y =$

$f(x_1, x_2, \dots, x_n)$ com base nas amostras fornecidas. A classificação difere desse cenário, pois a classe é categórica em vez de numérica (TORGO; GAMA, 1996).

- **Aprendizado não supervisionado:** ao contrário do aprendizado supervisionado mencionado anteriormente, não há respostas corretas e não há um professor. Os algoritmos são deixados por conta própria para descobrir e apresentar a estrutura interessante nos dados. Os algoritmos de aprendizado não supervisionado aprendem algumas características a partir dos dados. Quando novos dados são introduzidos, eles usam as características previamente aprendidas para reconhecer a classe dos dados. Isso é principalmente usado para agrupamento e redução de características (MAHESH; 2020).

K-Means e PCA (Análise de Componentes Principais) representam dois dos algoritmos de aprendizado não supervisionado mais populares e amplamente utilizados na atualidade. O PCA é um método usado para reduzir o número de variáveis em seus dados, extraindo uma variável importante de um grande conjunto de dados (JOLLIFFE, 2022).

Ele reduz a dimensão dos dados com o objetivo de reter o máximo de informação possível. Já O K-Means é a técnica de aprendizagem não-supervisionada mais simples, e consiste em fixar k centroides (de maneira aleatória), sendo um para cada cluster (PAN; HAN; YANG, 2010).

Em outras palavras, ele busca agrupar dados com características em comum envolta desses centróides.

- **Por Reforço:** É aquele em que o aprendizado se dá por meio de recompensas ou não ao indutor, o que vai depender do seu desempenho em aproximar a função desejada.

Sutton e Barto (1998) descrevem a aprendizagem por reforço (AR) como uma técnica de Inteligência Artificial que capacita um agente a adquirir conhecimento por meio de sua interação com o ambiente em que opera. Essa aprendizagem é baseada na compreensão do estado atual do agente no ambiente, nas ações que ele executa nesse ambiente e nas mudanças resultantes dessas ações. Em termos formais, o AR utiliza uma estrutura que inclui estados, ações e recompensas (GUELPELI et al., 2003).

O sucesso do ML nos tempos recentes foi inicialmente marcado por melhorias significativas em algumas tecnologias existentes, como no campo do reconhecimento de imagens. Em grande medida, esses avanços constituíram as primeiras demonstrações do impacto que os métodos de ML podem ter em tarefas especializadas. Mais recentemente, aplicativos tradicionalmente inacessíveis

a softwares automatizados foram habilitados com sucesso, em particular, pela tecnologia de Deep Learning (DL) (CARLEO et al., 2019).

2.3.2 Deep Learning

O DL é um tipo de ML que envolve o uso de redes neurais com múltiplas camadas para aprender representações de dados com vários níveis de abstração. Esses métodos melhoraram dramaticamente o que há de mais moderno em reconhecimento de fala, reconhecimento visual de objetos, detecção de objetos e muitos outros domínios, como descoberta de medicamentos e genômica. O DL descobre estruturas complexas em grandes conjuntos de dados usando o algoritmo de retropropagação para indicar como uma máquina deve alterar seus parâmetros internos que são usados para calcular a representação em cada camada a partir da representação na camada anterior (LECUN et al., 2015).

De acordo com HAYKIN, 2001 a rede neural se assemelha ao cérebro humano em dois aspectos básicos: a) o conhecimento é adquirido pela rede a partir de seu ambiente, por intermédio do processo de aprendizagem; b) forças de conexão entre neurônios (pesos sinápticos) são utilizadas para armazenar o conhecimento adquirido.

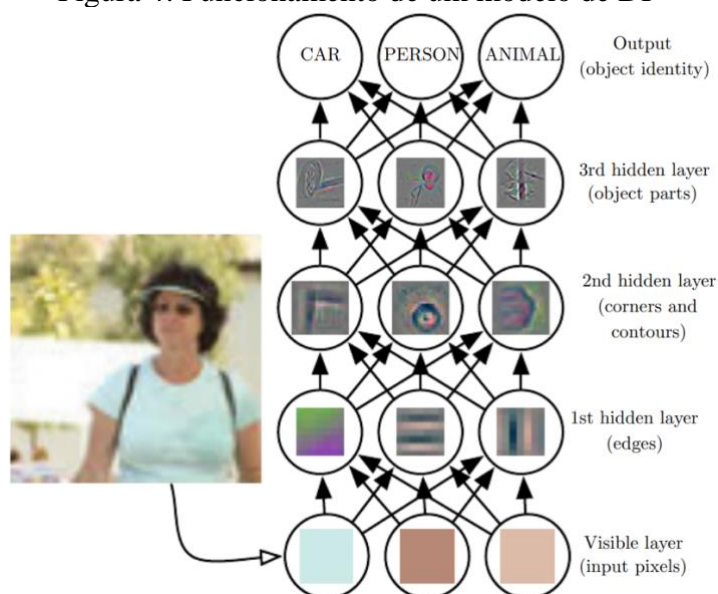
As tradicionais distinções entre aprendizado tradicional e profundo continuam a ser debatidas. No entanto, fundamental para o aprendizado profundo é o conceito de Rede Neural Artificial (ANN). ANNs são modelos de aprendizado de máquina projetados para emular o cérebro humano e são caracterizados por uma ou mais camadas de nós interconectados (neurônios) que geram representações não lineares das características de entrada, sendo úteis para problemas supervisionados, não supervisionados ou semi-supervisionados (AIZENSTEIN; MOORE, 2023).

A maior parte dos modelos de DL corresponde simplesmente às ANNs com múltiplas camadas de nós interconectados. No contexto da aprendizagem supervisionada, ajustar modelos de ANNs com mais camadas (i.e., modelos mais profundos) tem o potencial de levar a uma melhor precisão de previsão, e esta é uma das principais razões pelas quais a aprendizagem profunda tem recebido considerável atenção recentemente (AIZENSTEIN; MOORE, 2023).

Figura 4 mostra como um modelo de aprendizado profundo funciona. Quando um computador vê uma imagem, como esta com todos os seus pixels, entender o que está nela é difícil. O modelo de DP resolve isso dividindo a tarefa em partes mais simples. Primeiro, a imagem é apresentada à camada inicial, onde são observados os pixels. Depois, uma série de camadas "ocultas" analisa a imagem, procurando por características como bordas, cantos e partes de objetos.

Cada camada oculta procura padrões um pouco mais complexos do que a anterior. Por fim, o modelo usa essas informações para identificar o que está na imagem.

Figura 4: Funcionamento de um modelo de DP



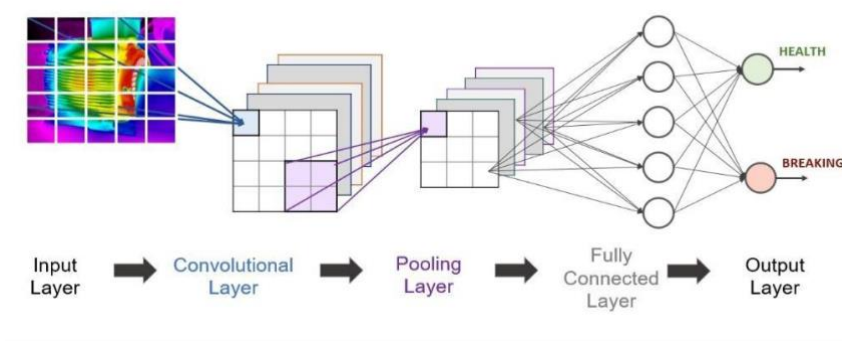
Fonte: HEATON, 2018

2.3.3 CNN

Uma rede neural convolucional (CNN) é um tipo de algoritmo de aprendizado profundo que tem se tornado popular especialmente para classificação de imagens, devido à capacidade desse algoritmo de encontrar padrões em imagens (LIN; YUH, 2022). Uma das primeiras aplicações bem-sucedidas das CNNs foi na classificação de imagens. Em particular, o artigo seminal de KRIZHEVSKY, 2012 demonstrou que uma CNN profunda (chamada AlexNet) poderia superar significativamente o estado-da-arte em um conjunto de dados de classificação de imagens de grande escala (ImageNet). A Figura 5 demonstra o processo de inferência de uma CNN, desde a entrada do dado da imagem como uma matriz até a predição.

As CNNs são uma classe de modelos de aprendizado profundo especialmente adequados para processamento de dados de entrada com estrutura de grade, como imagens, sinais de áudio e dados de séries temporais. As CNNs são baseadas em camadas convolucionais, que usam operações de convolução para extrair recursos relevantes dos dados de entrada. De acordo com LECUN, 2015, atualmente uma típica arquitetura de uma Rede Convolucional é dividida em uma série de estágios. Os primeiros estágios são compostos de dois tipos de camadas, as camadas de convolução e as camadas de pooling.

Figura 5: Exemplo de uma imagem de uma rede neural convolucional em funcionamento.



Fonte: MAIOR et al., 2021

CNNs demonstraram excelente desempenho em tarefas como classificação de dígitos escritos à mão e detecção de faces (ZEILER; FERGUS, 2014). Vários fatores são responsáveis por esse interesse renovado em modelos de convnet: (i) a disponibilidade de muitos conjuntos de treinamento maiores, com milhões de exemplos rotulados; (ii) poderosas implementações de GPU, tornando a prática de treinamento de modelos muito grandes e (iii) melhores estratégias de regularização de modelos, como Dropout (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

Nos últimos anos, o avanço na visão computacional com redes neurais convolucionais CNN despertou muito interesse em pesquisas. Uma das abordagens é a representação de sinais acústicos (fala, música ou som) em multi-resolução tempo-frequência (T-F) como entrada para a CNN. Uma camada convolucional da CNN processa uma imagem da fala ou som (espectrograma ou qualquer outra representação T-F) (CHAURASIYA, 2020).

Clique ou toque aqui para inserir o texto. A principal camada das CNNs é a convolução, cuja função é aplicar máscaras em dados de entrada, podendo ser não apenas imagens, mas também sequências de texto, com base em uma vizinhança de pixels. A saída produzida nessa operação são os filtros de convolução (matrizes) que armazenam os pesos das conexões entre os neurônios (PARKHI et al., 2015).

A camada de convolução consiste em mapas de atributos, conectados a cada unidade da camada anterior através de um conjunto de parâmetros compartilhados entre todas as unidades e pode possuir ReLUs (Rectified Linear Units), neurônios com função de ativação mais utilizada, definida como a não-linearidade aplicados na saída de cada camada convolucional (KRIZHEVSKY; SUTSKEVER; HINTON, 2012).

$$R(z) = \max(0, z)$$

A ativação é uma estrutura muito importante na camada de convolução. Essa realiza um ajuste entre um conjunto de neurônios, com uma função de ativação sobre uma convolução. Isso produz os mapas de características que armazenam as informações aprendidas dos filtros (PARKHI et al., 2015).

Existem outras funções de ativação tais como: sigmóides e tangente hiperbólica, mas ReLU é a mais utilizada para CNNs. Além disso, essa função de ativação reduz o tempo de convergência dos parâmetros, pois ela é simplesmente a função identidade para valores positivos (KRIZHEVSKY; SUTSKEVER; HINTON, 2012). A Figura 6 mostra a aplicação da função ReLU em uma matriz.

Figura 6 – Aplicação da função ReLU.



Fonte: Adaptado de HIJAZI et al., 2015

A camada de pooling é responsável por reduzir a dimensionalidade dos mapas de características, diminuindo a largura e a altura dos mapas de características. A operação de pooling possibilita uma invariância espacial. O agrupamento de características, na maioria das arquiteturas de convolução, utilizam uma função de Max pooling. Essa determina o valor máximo de agrupamento em uma vizinhança retangular. (ZHOU; CHELLAPPA, 1992)

As próximas camadas das CNNs desempenham o papel de regressão das ativações. Em qualquer rede desse tipo, após a camada de pooling, é necessário ao menos uma camada totalmente conectada (Fully Connected Layer- FCL). Elas servem para criar caminhos de decisões, a partir dos filtros obtidos na camada anterior (PARKHI et al., 2015).

A última camada, a FCL, por sua vez, é responsável por realizar a classificação dos dados. Nessa situação, uma função determina a identificação das saídas em classes. A função mais utilizada é a Softmax (para problemas de múltiplas classes) ou Sigmoid (problemas binários) (PARKHI et al., 2015).

O treinamento das RNCs é realizado na maioria dos casos com o backpropagation, que ajusta os pesos w dos neurônios pelo erro mensurado entre a verdade e a predição da rede, utilizando os componentes do vetor gradiente.

3 CONJUNTO DE DADOS

O conjunto de dados utilizado nesta pesquisa foi o IEEE PHM 2012 Data Challenge. O desafio teve foco na estimação do RUL dos rolamentos. Os conjuntos de dados do desafio foram fornecidos pelo Instituto FEMTO-ST (Besançon - França, [//www.femto-st.fr/](http://www.femto-st.fr/)). Os experimentos foram realizados em uma plataforma experimental de laboratório (PRONOSTIA) que permite a degradação acelerada dos rolamentos sob constante e/ou variável condições de operação, enquanto coleta dados de monitoramento de saúde online (velocidade de rotação, força de carga, temperatura, vibração) (NECTOUX et al., 2012).

No desafio, o conjunto de dados contém um conjunto de treino de 6 rolamentos que estão operando em 3 diferentes condições, e um conjunto de teste com mais 11. Existem no total 17 conjuntos de dados obtidos em três condições de operação diferentes. Todos os rolamentos são operados desde as condições de teste até a falha, uma vez que o tempo de vida útil restante só pode ser medido quando o rolamento apresenta uma falha. A Tabela 1 ilustra essa divisão entre o conjunto de treinamento e teste.

Tabela 1: Conjuntos de treino e teste

	Condições de operação		
	Condição 1	Condição 2	Condição 3
Conjunto de treinamento	Rolamento1_1	Rolamento2_1	Rolamento3_1
	Rolamento1_2	Rolamento2_2	Rolamento3_2
Conjunto de teste	Rolamento1_3	Rolamento2_3	Rolamento3_3
	Rolamento1_4	Rolamento2_4	
	Rolamento1_5	Rolamento2_5	
	Rolamento1_6	Rolamento2_6	
	Rolamento1_7	Rolamento2_7	

Fonte: Dataset do IEEE 2012 PHM Prognostic Challenge

Tanto os conjuntos de dados de treinamento quanto de teste foram fornecidos em pastas compactadas no formato "7z". Os parâmetros de aquisição de dados estão listados abaixo:

- Sinais de vibração (horizontal e vertical);
 - Frequência de amostragem: 25,6 kHz;

- Gravações: São registradas 2560 amostras (ou seja, 1/10 s) a cada 10 segundos.;
- Sinais de temperatura;
 - Frequência de amostragem: 10 Hz;
 - Gravações: São registradas 600 amostras a cada minuto;

Para cada arquivo ASCII, os dados foram organizados conforme ilustrado abaixo na Tabela 2

Tabela 2: Informações sobre os tipos do conjunto de dados

Sinais	1	2	3	4	5	6
Vibração	Hour	Minute	Second	μ -second	Horiz. accel	vert. accel.
Temperatura	Hour	Minute	Second	0.x second	Rtd sensor	

Fonte: Dataset do IEEE 2012 PHM Prognostic Challenge

Nesse trabalho a proposta foi um pouco diferente. Decidimos por nos concentrar apenas na condição 1. Os conjuntos de treino e teste foram divididos de uma forma diferente também. Para o primeiro caso foi utilizado os rolamentos 1, 2, 3, 4, 5 e 6 para o treinamento da rede e o conjunto 7 para teste como ilustrado pela Tabela 3.

Tabela 3: Conjuntos de treino e teste para o projeto primeiro caso

	Rolamentos na condição de operação 1
Conjunto de treinamento	Rolamento1_2
	Rolamento1_3
	Rolamento1_4
	Rolamento1_5
	Rolamento1_6
	Rolamento1_7
Conjunto de teste	Rolamento1_1

Fonte: Autor

Para o caso 2, foi utilizado os rolamentos 1, 2, 3, 4, 5 e 7 para o treinamento e o conjunto 6 para o teste como ilustrado na tabela 4.

Tabela 3: Conjuntos de treino e teste para o projeto segundo caso

Conjunto de treinamento	Rolamentos na condição de operação 1
	Rolamento1_1
	Rolamento1_2
	Rolamento1_3
	Rolamento1_4
	Rolamento1_5
	Rolamento1_7
Conjunto de teste	Rolamento1_6

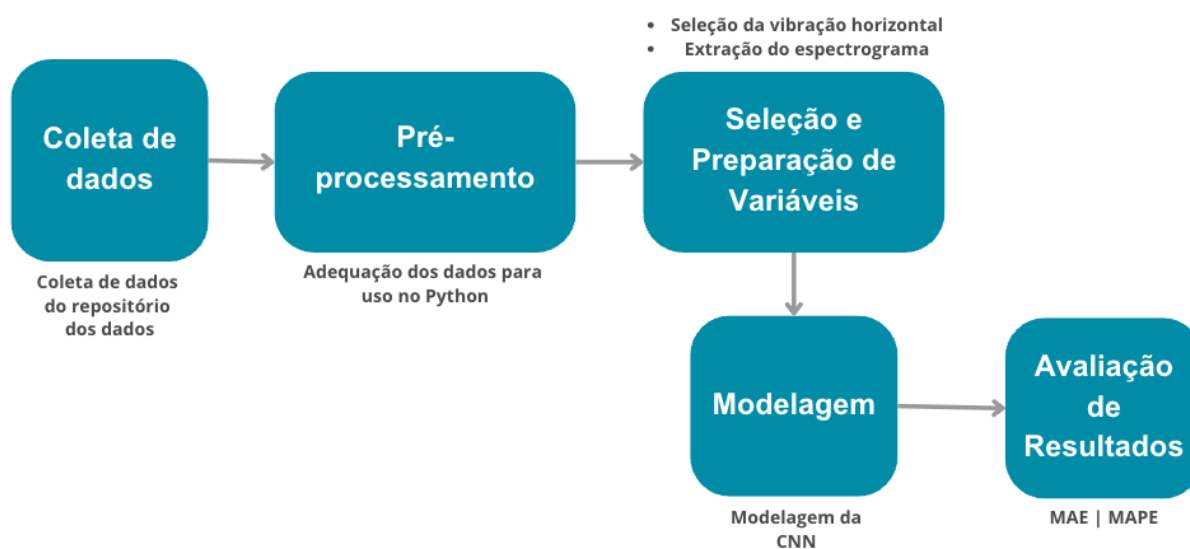
Fonte: Autor

Como visto, nesse conjunto de dados temos informações de vibração e temperatura ao longo do tempo. O esforço foi concentrado em lidar somente com os dados de vibração, e utilizando somente os valores da vibração horizontal. O espectrograma associado a essa variável foi extraído e preparado para ser utilizado como input inicial da construção do modelo.

4 METODOLOGIA

A metodologia aplicada é apresentada abaixo. Dados de vibração de rolamentos são usados como insumo para a metodologia. A Figura 7 demonstra os passos da metodologia.

Figura 7 – Ilustração da metodologia



Fonte: Autor

4.1 Coleta de Dados

Nesta fase, foi implementado um script de download para adquirir os dados necessários e armazená-los localmente em um diretório denominado "data/".

4.2 Pré-processamento de Dados

O conjunto de dados passou por um processo de pré-processamento para otimizar sua usabilidade no ambiente Python, utilizando a biblioteca Pandas. Foram realizadas transformações específicas para garantir a adequação dos dados ao escopo da análise. Os dados foram fornecidos em vários arquivos .csv, então foi feito um tratamento para agrupar esses dados e utilizá-los da melhor forma.

4.3 Seleção e Preparação de Variáveis

Na etapa subsequente, focamos na variável de interesse, que neste caso é a "frequência horizontal". O espectrograma associado a essa variável foi extraído e preparado para ser utilizado como input na construção do modelo.

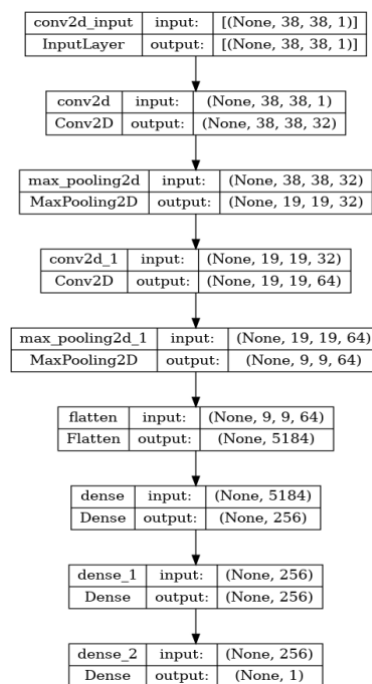
4.4 Modelagem

O modelo adotado para este estudo é uma CNN. A implementação foi realizada utilizando a biblioteca TensorFlow, uma ferramenta amplamente reconhecida para o desenvolvimento de redes neurais. A escolha da CNN baseou-se em sua capacidade de lidar eficazmente com tarefas relacionadas à imagem e processamento de dados multidimensionais.

A concepção das camadas tem como objetivo a criação de imagens adicionais, as quais podem ser consideradas como matrizes, uma vez que no âmbito do nosso problema não estamos lidando diretamente com imagens convencionais. O propósito dessa abordagem é gerar representações de dados de dimensões progressivamente menores à medida que as camadas são empilhadas.

Consequentemente, observa-se que ao progredir de uma camada para a próxima, são geradas imagens adicionais com resoluções reduzidas, o que viabiliza a capacidade de identificar padrões com maior eficiência e precisão. A figura 8 ilustra as camadas da rede utilizada.

Figura 8 – Camadas da CNN utilizada.



Fonte: Autor

Primeira Camada de Convolução:

- Filtros: 32
- Tamanho do Kernel: 5x5
- Função de Ativação: ReLU (Rectified Linear Unit)
- Camada de Entrada: A forma da camada de entrada é determinada pela forma dos dados de treinamento (variável train).
- Propósito: Esta camada realiza a convolução dos filtros de tamanho 5x5 com os dados de entrada, aplicando a função de ativação ReLU. Ela é responsável por extrair características de baixo nível das imagens de entrada.
- Input: Nessa camada temos como input as imagens geradas pelo espectrograma da vibração horizontal.

Camada de MaxPooling (Pooling Máximo) 1:

- Pooling Máximo: Tamanho do pool 2x2 com um passo (stride) de 2.
- Propósito: Esta camada reduz as dimensões espaciais das características extraídas pela camada de convolução anterior. Ela ajuda a preservar as características mais importantes e reduz a quantidade de parâmetros na rede.

Segunda Camada de Convolução:

- Filtros: 64
- Tamanho do Kernel: 5x5
- Função de Ativação: ReLU
- Camada de Entrada: Assume a forma da saída da primeira camada de convolução.
- Propósito: Assim como a primeira camada de convolução, esta camada realiza a convolução com filtros maiores (5x5) para extrair características de mais alto nível das imagens. A função de ativação ReLU é aplicada novamente.

Camada de MaxPooling (Pooling Máximo) 2:

- Pooling Máximo: Tamanho do pool 2x2 com um passo (stride) de 2.
- Propósito: Mais uma vez, esta camada reduz as dimensões espaciais das características extraídas pela segunda camada de convolução. Isso ajuda a preparar os dados para as camadas totalmente conectadas.

Camada de Flattening:

- Propósito: Esta camada transforma as características em um vetor unidimensional. Isso é necessário para conectar as camadas totalmente conectadas.

Primeira Camada Totalmente Conectada (Dense):

- Neurônios: 256
- Função de Ativação: ReLU
- Propósito: Camada totalmente conectada que realiza operações lineares e aplica a função de ativação ReLU. Ela aprende relações mais complexas nos dados.

Segunda Camada Totalmente Conectada (Dense):

- Neurônios: 256
- Função de Ativação: ReLU
- Propósito: Outra camada totalmente conectada com 256 neurônios e função de ativação ReLU.

Camada de Saída (Dense):

- Neurônios: 1
- Função de Ativação: Linear
- Propósito: Esta é a camada de saída da rede.

A função softmax não é tipicamente usada nas camadas de saída de Convolutional Neural Networks (CNNs) para tarefas de classificação de imagens. A função softmax é usada para converter um vetor de valores reais em uma distribuição de probabilidade, onde a soma de todas as saídas é igual a 1. Isso é útil quando você tem várias classes e deseja atribuir uma probabilidade para cada classe.

A ativação 'Linear' na última camada da rede mostrada é devido ao nosso problema buscar uma resposta numérica, ou seja, uma tarefa que a rede está configurada para uma tarefa de regressão, onde a saída é um valor numérico. Abaixo é mostrado uma ilustração de como ficou nosso modelo e suas camadas.

4.5 Avaliação de Resultados

A análise dos resultados foi conduzida com base em métricas de desempenho, especificamente a métrica de mean absolute error (MAE) e mean absolute percentage error (MAPE). Essas métricas, entre outras são comumente utilizadas para avaliar a qualidade de modelos

de regressão, fornecendo uma medida quantitativa do desempenho do modelo em relação aos dados observados. Essas métricas serão utilizadas para fazer comparações com a literatura mais a frente.

A fórmula convencional do MAE e MAPE usada também avaliando a função objetiva, geralmente são definidas abaixo (ZHANG; NIU, 2024).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \bar{y}_i|$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \bar{y}_i|}{y_i}$$

4.6 Comparação com a literatura

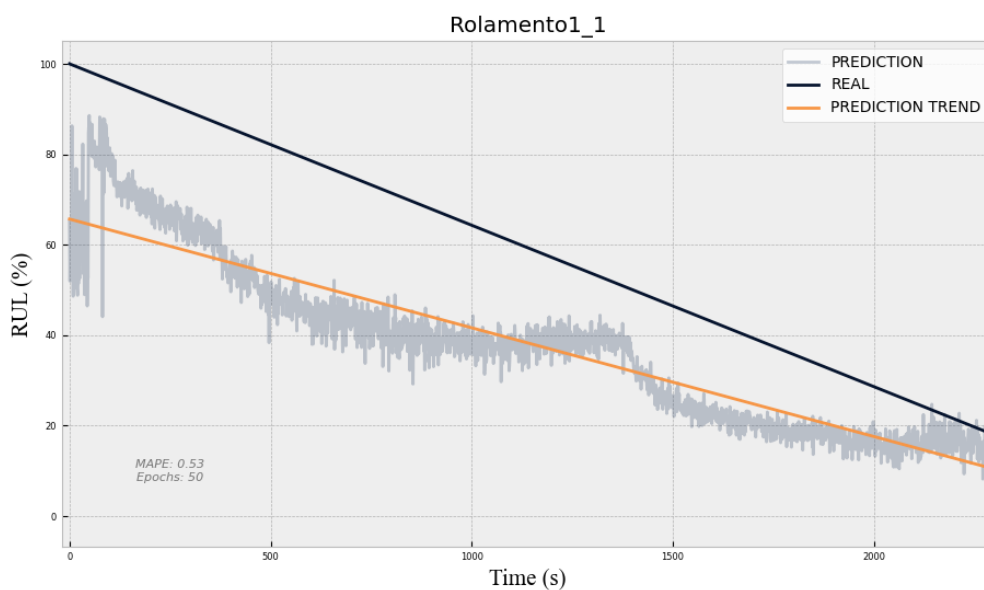
Na ultima etapa foi realizada uma comparação com a literatura para identificar a performance de cada abordagem.

5 RESULTADOS

Para o caso 1, onde foi considerado os rolamentos 2, 3, 4, 5, 6 e 7 para treino e o rolamento 1 para teste foi obtido um MAE de 5095.13. Esse valor não traz muita informação sobre se essa variação é grande ou não. Conjuntamente com o MAE, foi calculado o MAPE, que trás a informação percentual do MAE. Foi obtido um valor 53% de variação para o MAPE. Comparando com a literatura, o artigo de LI, 2019 trás uma previsão de MAPE para o mesmo rolamento 1 na condição 1 de 62,3%.

O artigo traz um comparação entre diferentes abordagens para prever o RUL para cada rolamento descrito no conjunto de dados utilizado neste trabalho. O modelo de DL implementado foi uma rede neural profunda convencional, também conhecida como multi-layer perceptron (MLP). A Figura 9 ilustra o RUL real para o rolamento 1 na condição 1 e o RUL predito para esse rolamento, junto com uma reta de tendência para a predição do modelo.

Figura 9 – RUL real vs RUL predito para o rolamento 1 na condição 1.



Fonte: Autor

É esperado que o erro nos conjuntos de treino sejam menores pois estamos fazendo uma predição para dados já vistos pelo modelo e utilizados justamente no seu treino. Logo, é esperado que ele tenha apreendido o comportamento de dados já vistos.

Tabela 4: Valores do MAE e MAPE para o treino e teste do rolamento 1 na condição 1.

Conjunto de dados	MAE	MAPE
Treino	1410.01	0.33
Teste	5095.13	0.53

Fonte: Autor.

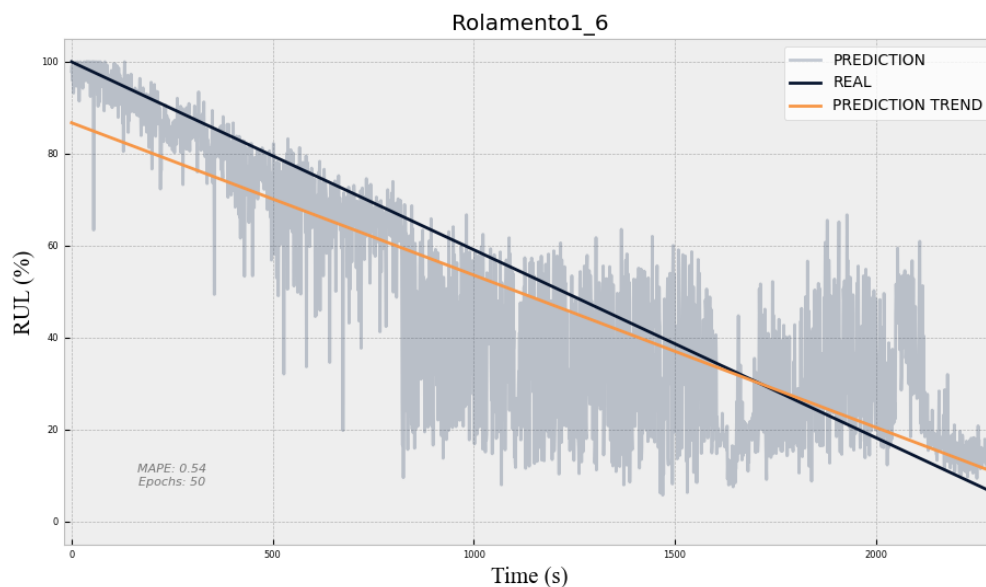
Tabela 5: Valores de MAPE para rolamento 1 na condição 1 para o modelo de CNN Utilizado nesse trabalho e o MLP utilizado em LI; ZHANG; DING, 2019

Conjunto de dados	MAPE
CNN	0.53
MLP	0.62

Fonte: Autor..

Para o caso 2, onde foi considerado os rolamentos 1, 2, 3, 4, 5 e 7 para treino e o rolamento 6 para teste foi obtido um MAE e MAPE de 15% e 53% respectivamente. Comparando com resultado obtido por LI, 2019 foi de 49,4% de MAPE. A Figura 10 ilustra o RUL real para o rolamento 6 na condição 1 e o RUL predito para esse rolamento.

Figura 10 – RUL real vs RUL predito para o rolamento 6 na condição 1.



Fonte: Autor

Tabela 6: Valores do MAE e MAPE para o treino e teste do rolamento 6 na condição 1.

Conjunto de dados	MAE	MAPE
Treino	1073.77	0.15
Teste	2762.17	0.536

Fonte: Autor.

Tabela 7: Valores de MAPE para rolamento 6 na condição 1 para o modelo de CNN Utilizado nesse trabalho e o MLP utilizado em LI; ZHANG; DING, 2019

Conjunto de dados	MAPE
CNN	0.53
MLP	0.49

Fonte: Autor.

6 CONCLUSÃO

Nesse trabalho é proposto uma proposta para a predição do RUL utilizando uma CNN. Comparando resultados com a literatura foi visto que para o caso 1 essa abordagem teve um desempenho melhor, para o caso 2 a abordagem de MLP foi mais bem-sucedida. Pode-se perceber que para a abordagem 1, com os devidos conjunto de treino e teste a proposta de arquitetura feita nesse trabalho conseguiu uma performance melhor. Se adaptando e aprendendo melhor o comportamento desse conjunto de dados. Para o caso 2 isso não ocorreu, o conjunto de treino e teste para esse caso 2 são diferentes, logo a abordagem proposta não conseguiu tirar melhor proveito dessa situação para preformar melhor a literatura. De forma geral podemos perceber que para diferentes rolamentos uma abordagem ou outra pode ser mais adequada. Além de pontos como otimização de parâmetros nesse trabalho não foram considerados. Os resultados experimentais sugerem que o método proposto é promissor para problemas prognósticos e é adequado para aplicações industriais.

Apesar dos bons resultados prognósticos alcançados pelo método proposto, deve-se destacar que são necessários dados de treinamento rotulados suficientes. Isso é geralmente exigido por todas as abordagens baseadas em dados, especialmente para problemas prognósticos.

Como próximos passos, se destacam pontos como realizar ajustes na arquitetura da CNN, como modificar o número de camadas, utilização de mais variáveis mudando a abordagem, o tamanho dos filtros, a função de ativação, ou adicionar técnicas como dropout e batch normalization para melhorar o desempenho do modelo. Utilizar conjuntos de dados de rolamentos de diferentes fabricantes, condições operacionais e tipos de falhas, para avaliar a generalização do método proposto e entender sua robustez em diferentes contextos. Experimentar diferentes arquiteturas de redes neurais, como Redes Neurais Recorrentes (RNNs).

Como output desse trabalho temos esse repositório no github https://github.com/marcosvliras/CNN_Bearing onde descreve como é possível replicar o projeto.

REFERÊNCIAS BIBLIOGRÁFICAS

AIZENSTEIN, H.; MOORE, R. C. Deep Learning and Geriatric Mental Health ARTICLE IN PRESS. 2023.

ALAN KARDEC PINTO; JULIO NASCIF. **MANUTENÇÃO: FUNÇÃO ESTRATÉGICA - 2ªED.** 2. ed. [s.l: s.n.].

ARMANDO DE ALMEIDA MURÇA, V. et al. Aplicação da filosofia Lean na área da Manutenção. 2012.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. ABNT. NBR 5462: 1994. Confiabilidade e Manutenibilidade. Rio de Janeiro: ABNT, 1994. , 1994.

CAI, B. et al. A novel RUL prognosis methodology of multilevel system with cascading failure: Subsea oil and gas transportation systems as a case study. **Ocean Engineering**, v. 242, 15 dez. 2021.

CARLEO, G. et al. Machine learning and the physical sciences. **APSG Carleo, I Cirac, K Cranmer, L Daudet, M Schuld, N Tishby, L Vogt-Maranto, L ZdeborováReviews of Modern Physics, 2019•APS**, v. 91, n. 4, 6 dez. 2019.

CHAUASIYA, H. ScienceDirect-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>) Peer-review under responsibility of the scientific Time-Frequency Representations: Spectrogram, Cochleogram and Correlogram-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>) Peer-review under responsibility of the scientific committee of the International Conference on Computational Intelligence and Data Science (ICCIDS 2019). **Procedia Computer Science**, v. 167, p. 0–000, 2020.

CHEN, Z. et al. Mechanical fault diagnosis using convolutional neural networks and extreme learning machine. **Elsevier**, 2019.

COSTA, Y. Reconhecimento de gêneros musicais utilizando espectrogramas com combinação de classificadores. 2013.

CROCKETT, B. et al. Capturing ultra-broadband complex-fields of arbitrary duration using a real-time spectrogram. **pubs.aip.org**, 2023.

DJP AUTOMAÇÃO. **Sensores industriais – Descubra sua importância para a indústria moderna.**

GUELPELI, M. et al. Utilização de aprendizagem por reforço para modelagem autônoma do aprendiz em um tutor inteligente. **ic.uff.br**, 2003.

HAYKIN, S. **Redes neurais: princípios e prática**. [s.l: s.n.].

HEATON, J. Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning. **Genetic Programming and Evolvable Machines**, v. 19, n. 1–2, p. 305–307, jun. 2018.

HIJAZI, S. et al. Using convolutional neural networks for image recognition. **site.eet-china.com**, 2015.

HOYLAND, A.; RAUSAND, M. **System reliability theory: models and statistical methods**. [s.l: s.n.].

JANA, G. et al. A 1D-CNN-spectrogram based approach for seizure detection from EEG signal. **Elsevier**, 2020.

JOLLIFFE, I. A 50-year personal journey through time with principal component analysis. **Journal of Multivariate Analysis**, v. 188, p. 104820, 1 mar. 2022.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. 2012.

LASI, H. et al. Industry 4.0. **Springer**, v. 6, n. 4, p. 239–242, 1 ago. 2014.

LECUN, Y. et al. Deep learning. **nature.com**, 2015.

LEON COHEN. **Time-frequency Analysis**. [s.l.] Prentice Hall PTR, 1995.

LI, X.; ZHANG, W.; DING, Q. Deep learning-based remaining useful life estimation of bearings using multi-scale feature extraction. **Reliability Engineering and System Safety**, v. 182, 2019.

LIAKOS, K. et al. Machine learning in agriculture: A review. **mdpi.com**, 2018.

LIN, E.; YUH, E. L. Computational Approaches for Acute Traumatic Brain Injury Image Recognition. **Frontiers in Neurology**, v. 13, 9 mar. 2022.

LUCIFREDI, A. et al. Application of multiregressive linear models, dynamic kriging models and neural network models to predictive maintenance of hydroelectric power systems. **Elsevier**, 2000.

MAHESH, B.-I. J. OF S. AND R.; 2020. Machine learning algorithms-a review. **researchgate.netB MaheshInternational Journal of Science and Research (IJSR).[Internet], 2020•researchgate.net**, 2018.

MAIOR, C. B. S. et al. Convolutional neural network model based on radiological images to support COVID-19 diagnosis: Evaluating database biases. **PLOS ONE**, v. 16, n. 3, p. e0247839, 1 mar. 2021.

MIRSHAWKA, V.; ... N. O. CUSTOS DA NÃO EFICÁCIA: A VEZ; 1993, UNDEFINED. Manutenção-combate aos custos da não eficácia: a vez do Brasil. **pesquisa.bvsalud.org**, 1993.

MOHANTY, A. **Machinery condition monitoring: Principles and practices**. [s.l: s.n.].

NECTOUX, P. et al. PRONOSTIA: An experimental platform for bearings accelerated degradation tests. **hal.science**, p. 1–8, 2012.

PAN, J. YANG; HAN, J.; YANG, S. E. A neural network based method for detection of weak underwater signals. **Journal of Marine Science and Application**, v. 9, n. 3, p. 256–261, 2010.

PARKHI, O. et al. Deep face recognition. **ora.ox.ac.uk** O Parkhi, A Vedaldi, A Zisserman **BMVC 2015-Proceedings of the British Machine Vision Conference 2015**, 2015•**ora.ox.ac.uk**, 2015.

PERES, R. et al. Industrial artificial intelligence in industry 4.0-systematic review, challenges and outlook. **ieeexplore.ieee.org** RS Peres, X Jia, J Lee, K Sun, AW Colombo, J Barata **IEEE Access**, 2020•**ieeexplore.ieee.org**, 2020.

RASCHKA, S. **Python machine learning**. [s.l: s.n.].

SAHA, D. et al. Development of intelligent fault diagnosis technique of rotary machine element bearing: A machine learning approach. **mdpi.com** DK Saha, ME Hoque, H Badihi **Sensors**, 2022•**mdpi.com**, 2022.

SAN MARTIN, G. et al. Deep variational auto-encoders: A promising tool for dimensionality reduction and ball bearing elements fault diagnosis. **Structural Health Monitoring**, v. 18, n. 4, p. 1092–1128, 1 jul. 2019.

SAXENA, A. et al. Metrics for offline evaluation of prognostic performance. **books.google.com** A Saxena, J Celaya, B Saha, S Saha, K Goebel **International Journal of Prognostics and health management**, 2010•**books.google.com**, v. 1, 2010.

SCAIFE, A. D. Improve predictive maintenance through the application of artificial intelligence: A systematic review. **Results in Engineering**, v. 21, p. 101645, mar. 2024.

SI, X. et al. Remaining useful life estimation—a review on the statistical data driven approaches. **Elsevier**, 2011a.

SI, X. et al. Remaining useful life estimation—a review on the statistical data driven approaches. **Elsevier**, 2011b.

SOARES, L. DE O. et al. Aplicação de Machine Learning e Descritores Estatísticos para Diagnóstico de Falhas em Rolamentos de Motores Elétricos Industriais. **sba.org.br** L de Oliveira Soares, LA Pinto, DA Lobao **Congresso Brasileiro de Automática-CBA, 2020•sba.org.br**, 2020.

SOUTO MAIOR, C. BEZERRA. **REMAINING USEFUL LIFE PREDICTION VIA EMPIRICAL MODE DECOMPOSITION, WAVELETS AND SUPPORT VECTOR MACHINE**. Recife: [s.n.].

SUTRISNO, E. et al. Estimation of remaining useful life of ball bearings using data driven methodologies. **ieeexplore.ieee.org**, 2012.

TANDON, N.; INTERNATIONAL, A. C.-T.; 1999, UNDEFINED. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. **Elsevier**, 1999.

TORGO, L.; GAMA, J. Regression by classification. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 1159, p. 51–60, 1996.

ZAWARESUMEDHA. Plotting a Spectrogram using Python and Matplotlib. Disponível em: <<https://acervolima.com/tracando-um-espectrograma-usando-python-e-matplotlib/>>. Acesso em: 22 dez. 2023.

ZEILER, M. D.; FERGUS, R. Visualizing and understanding convolutional networks. **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 8689 LNCS, n. PART 1, p. 818–833, 2014.

ZHANG, D.; NIU, B. Leveraging online reviews for hotel demand forecasting: A deep learning approach. **Information Processing and Management**, v. 61, p. 306–4573, 2024.

ZHANG, H. et al. mixup: Beyond empirical risk minimization. **arxiv.org**, 2017.

ZHOU, Y.-T.; CHELLAPPA, R. Computational Neural Networks. p. 6–14, 1992.

ZU, S. et al. Dictionary learning based on dip patch selection training for random noise attenuation. **Geophysics**, v. 84, n. 3, p. V169–V183, 1 maio 2019.