**Otimização do monitoramento das condições de equipamentos de mineração, usando manutenção preditiva e IoT com Aprendizado de máquina**

Cid Clay Aparecido Quirino¹\*;Eder Costa Cassettari2

1 Cid Clay Quirino. Mestre em Engenharia Mecânica. Avenida Assis de Vasconcelos 542 – Campina; 66017-070 Belém, Para, Brasil

2 Nome da Empresa ou Instituição (opcional). Titulação ou função ou departamento. Endereço completo (pessoal ou profissional) – Bairro; 00000-000 Cidade, Estado, País

\*autor correspondente: nome@email.com

**Otimização do monitoramento das condições de equipamentos de mineração, usando manutenção preditiva e IoT com Aprendizado de máquina**

**Sumário Executivo**

O presente trabalho abordou a otimização de processos de manutenção preditiva em equipamentos de mineração utilizando tecnologias de IoT e aprendizado de máquina. A justificativa do estudo baseou-se na crescente demanda por maior confiabilidade e redução de falhas em ativos críticos de mineração, como escavadeiras hidráulicas, onde componentes menores muitas vezes não são monitorados adequadamente, sendo uma oportunidade de avançar neste processo. O objetivo foi promover o sensoriamento de pequenos componentes, permitindo o monitoramento contínuo de variáveis como temperatura e vibração, em contrapartida com o fluxo manual utilizado com rotina nas operações. A metodologia incluiu a coleta de dados em campo por sensores IoT instalados nos componentes e o uso de técnicas de aprendizado de máquina para a análise preditiva das falhas. A base de dados foi tratada e analisada utilizando algoritmos como RNN-GRU, RNN-LSTM, e SVR, com a projeção de quando os componentes atingiriam limites críticos de operação. Entre os principais resultados, observou-se que os modelos de aprendizado profundo, especialmente RNN-GRU e RNN-LSTM, apresentaram alto desempenho preditivo, reduzindo significativamente o risco de falhas não planejadas. O estudo concluiu que a integração de IoT com técnicas avançadas de modelagem pode otimizar a manutenção preditiva em ambientes de mineração, proporcionando ganhos operacionais e maior segurança.

**Palavras-chave**: manutenção preditiva; aprendizado de máquina; IoT; mineração; otimização de processos.

**Optimization of Mining Equipment Condition Monitoring Using Predictive Maintenance and IoT with Machine Learning**

**Abstract**

This study addressed the optimization of predictive maintenance processes for mining equipment by utilizing IoT technologies and machine learning. The motivation for the study stemmed from the growing demand for higher reliability and reduction of failures in critical mining assets, such as hydraulic excavators, where smaller components are often inadequately monitored, presenting an opportunity for advancement. The objective was to implement sensor-based monitoring of smaller components, enabling continuous tracking of variables such as temperature and vibration, compared to the manual monitoring processes commonly used in operations. The methodology involved field data collection through IoT sensors installed on components and the use of machine learning techniques for predictive failure analysis. The data were processed and analyzed using algorithms like RNN-GRU, RNN-LSTM, and SVR, projecting when components would reach critical operational limits. The main findings showed that deep learning models, particularly RNN-GRU and RNN-LSTM, demonstrated high predictive performance, significantly reducing the risk of unplanned failures. The study concluded that integrating IoT with advanced modeling techniques can optimize predictive maintenance in mining environments, offering operational gains and enhanced safety.

**Keywords:** predictive maintenance; machine learning; IoT; mining; process optimization

**Introdução**

Estudos recentes mostram avanços no monitoramento de equipamentos industriais, tendo como objetivos a redução a utilização de seres humanos devido aos riscos inerentes, dificuldade de acesso aos locais de execução, além da redução na parada dos equipamentos com a antecipação de falhas, desde a década de 90 temos este avanço como indica Tian, Y. M. et al. (1992) indicando o aumento da produtividade e utilização dos equipamentos industriais, ou seja, a eficiência e produção com o aumento no nível de automação estão cada vez mais elevados, colaborando ainda mais com a necessidade de melhorar o monitoramento da saúde dos ativos.

Acrescenta-se a isso a avaliação de que, os equipamentos industriais apresentam dificuldades mesmo com o avanço nas técnicas de automatização, como relatado por Schmidt and Berns (2013) , apontando o fato de a manutenção e inspeção de grandes estruturas com sistemas autônomos ainda é um problema sem solução, eles avaliaram que existem diversas possíveis abordagem com a utilização de robôs para automatizar as inspeções aumentar a qualidade e confiabilidade, contudo ainda existe muito espaço para avanço nestas técnicas em várias áreas.

Para Park C. et al (2016), a manutenção preditiva atrai mais interesse do que a manutenção de rotina, que é descrita como sendo a manutenção realizada quando ocorre uma falha na máquina. Eles atribuem isso ao fato de que as técnicas de manutenção preditiva ajudam a determinar a condição dos equipamentos ou sistemas em serviço, e possibilitam avaliar e prever quando a manutenção deve ser realizada. A manutenção preditiva permite o agendamento conveniente de ações corretivas e evita paradas inesperadas do equipamento, sendo que a chave é a informação certa no momento certo.

Ao saber quais equipamentos ou componentes precisam de manutenção, os responsáveis pelo departamento de manutenção podem antecipar as manutenções, e o que seriam paradas não planejadas, podem ser transformadas em paradas mais curtas e mais eficientes, aumentando assim a disponibilidade do equipamento. Esta abordagem geralmente utiliza técnicas estatísticas de controle de processos, para determinar em que ponto as futuras atividades de manutenção serão apropriadas.

Além disso, Gbadamosi et al (2021) observa que algumas abordagens atuais exigem o envio de inspetores para áreas, para realizar verificações de rotina, o que representa riscos para a saúde e a segurança dos trabalhadores, por outro lado, quando temos um monitoramento eficiente dos ativos com métodos inovadores de coleta online por exemplo, esse risco pode ser reduzido e até mesmo eliminado.

No contexto da mineração, alguns componentes, como motores e pneus, já possuem um monitoramento robusto. No entanto, componentes menores, como bombas e cilindros hidráulicos, são frequentemente monitorados de forma indireta, o que pode retardar a tomada de decisões em caso de variações. Skoczylas et al. (2023) sugerem que um sistema aprimorado de monitoramento é crucial para garantir a continuidade das operações e prever avarias, com o uso mais intensivo de técnicas voltadas a manutenção Preditiva.

Ferreira, B et al. (2022) aborda a demanda crescente por processos digitalizados a partir da evolução tecnológica na Indústria 4.0, a necessidade de acesso a dados de maneira mais rápida, intuitiva e barata, e indica pontos de desenvolvimento de soluções viáveis e de baixo custo para auxiliar na visualização de dados e na utilização destes em antecipação de falhas.

Artur Skoczylas et al. (2023) também avalia que na maioria das grandes empresas, as máquinas foram equipadas com sistemas que medem diversos parâmetros, como rotação do motor, temperatura do óleo, pressão do sistema hidráulico, temperatura dos componentes mais importantes, pressão dos pneus, consumo de combustível, velocidade de movimento etc. os sistemas podem enviar dados em tempo real, no entanto, isso não é praticado em grande escala no mercado, devido aos elevados custos operacionais dessa mudança.

Este potencial de melhora no monitoramento para abranger também os pequenos componentes, foi a oportunidade identificada neste estudo, que indica uma das possibilidades de avançar sobre o processo de Gerenciamento de Monitoramento de Condições (CMMS) ampliando a cobertura para os pequenos componentes, utilizando para tal, a aplicação de sensores de temperatura e vibração, possibilitando a coleta em tempo real de parâmetros ao longo de uma série temporal.

Com base nessa oportunidade, foi elaborado um projeto de sensoriamento com Internet das Coisas (IoT), para pequenos componentes em uma escavadeira hidráulica de mineração e aplicado em campo em uma situação real de operação de mina de céu aberto.

Por fim, este estudo foca no desenvolvimento e aplicação de alternativas viáveis para coleta, análise de projeção de saúde em pequenos componentes, tendo como principal objetivo avançar na manutenção preditiva e na redução de falhas prematura, menor tempo de inatividade, atendendo também os requisitos de segurança, com a menor exposição aos riscos humanos para a inspeção.

**Material e Métodos**

Para a análise dos dados deste estudo, será utilizada uma base de dados com 8226 observações de temperatura e vibração de ambas as bombas principais da 395, coletados no período de 14 dias de operação como amostra, posteriormente, os dados estarão no PipeLine online, e serão avaliados assim que forem coletados e disponibilizados.

Os dados foram extraídos em um formato .csv e importados do GitHub para o Python e bibliotecas a seguir:

**NumPy:** Biblioteca que fornece suporte para arrays multidimensionais, juntamente com a ampla coleção de funções matemáticas para operar com análise de dados, computação numérica.

**Satsmodels:** Biblioteca Python que oferece classes e funções para estimar e interpretar diversos modelos estatísticos, incluindo métodos para ajustar modelos de regressão, análise de séries temporais, testes estatísticos etc.

**Scikit-learn:** Biblioteca para aprendizado de máquina que fornece uma ampla variedade de algoritmos de aprendizado supervisionado e não supervisionado.

**PyGithub:** Biblioteca Python que fornece uma interface para interagir com a API do GitHub, facilitando a automatização de tarefas como criação de repositórios.

**TensorFlow:** Biblioteca de código aberto para aprendizado de máquina e inteligência artificial desenvolvida pelo Google, oferece uma estrutura tanto para construir quanto treinar modelos de aprendizado profundo, incluindo redes neurais convolucionais, redes neurais recorrentes e muito mais.

**Manutenção Preditiva, Preventiva e Corretiva**

A manutenção corretiva pode ser descrita,segundo Nascimento et al. (2020) como sendo a modalidade de manutenção aplicada após a ocorrência da falha, com a finalidade de colocar o ativo em funcionamento novamente ou quando há pera da funcionalidade do equipamento e se pretende reestabelecer a operação. Essa modalidade pode ser dividida em dois tipos, corretiva planejada e não planejada.

Já a manutenção preventiva, também de acordo com Nascimento et al. (2020) é a modalidade de manutenção aplicada com o intuito de diminuir a probabilidade de falha e manter o desempenho, é desenvolvida através de procedimentos, planos de manutenção e inspeções, executada em intervalos regulares, após os quais realiza-se a substituição ou o reparo de itens em que foram detectadas anomalias, podendo citar como sendo as atividades de manutenção preventiva.

Por último, e considerando sua importância na Industria 4.0, temos a manutenção preditiva, essa modalidade de manutenção se utiliza de dados para as análises em pontos específicos do ativo, tais como, conferência de nível de óleo, análise de vibração e análise de temperatura, sendo comumente aplicada com a finalidade de monitorar o padrão de funcionamento do ativo e identificar potencial de falha quando existente.

Esta última modalidade de manutenção, a preditiva, se caracteriza pelo uso de instrumentos de medição e coleta massiva de dados como ferramenta de apoio a decisão de uma intervenção. Neste sentido, o uso de IoTs vem sendo amplamente usado pela indústria, para o avanço das técnicas de preditiva, e será mais bem detalhado a seguir.

O Avanço do uso de IoTs e seu avanço na indústria 4.0, pode ser considerado a base deste estudo que visa otimizar a condição de equipamentos de mineração, como apresentado no tópico posterior, onde pode ser visto o uso da manutenção preditiva e IoT como sendo a base para aumentar a coleta e disponibilidade de dados de máquina, e avançar na atuação preditiva.

Alexandre R. Oliveira (2023) cita essa problemática em seu estudo, descrevendo de forma clara que as máquinas e equipamentos industriais não foram construídos para durar para sempre contudo, podem durar muito mais. Ele acrescenta que os eventos de falhas de ativos são processos de degradação, que podem ser didaticamente apresentados por meio da curva P-F visto na figura 1, e que busca representar a condição de um equipamento ou componente ao longo do tempo, dando claramente uma indicação sobre o avanço da condição x tempo, indicando a necessidade de agir proativamente para evitar a falha.

|  |
| --- |
| Tela de celular com texto preto sobre fundo branco  Descrição gerada automaticamente com confiança baixa |
| Figura 1. Intervalo P-F.  Fonte: Oliveira (2023). |

Oliveira (2023) em seu estudo, também aponta para utilização de tecnologias no monitoramento do intervalo P-F como sendo uma estratégia inteligente com maior assertividade e confiabilidade na manutenção dos seus equipamentos e com isso a equipe de manutenção pode evitar as quebras repentinas, e que as falhas identificadas entre os pontos P e F da curva, podem ser corrigidas evitando uma falha funcional com o uso de softwares, sensores e inteligência artificial realizando a coleta de dados das máquinas, para ter uma visão completa dos processos industriais.

**Internet das coisas IoT**

Como citado por Aguirre (2021), a IoT pode ser considerado um mundo onde os objetos físicos são perfeitamente à rede de informações se tornando parte ativa dos fluxos e processos de negócios Haller et al. (2015), este é um dos principais objetivos do estudo em questão, pois substitui em boa medida, a necessidade de um inspetor ir até o ativo expondo-se ao risco para coleta de dados que podem ser facilmente, e constantemente coletadas pelos sensores.

Aguirre (2021) comenta que o surgimento da IoT de baixo custo promete acesso generalizado a sensores e dados que podem ser usados para a tomada de decisões operacionais. Em seu estudo uso as IoT, foi elaborado em um caso real onde foi aplicado um sistema para informar mudanças na gestão operacional que resultaram na redução do tempo de carregamento, na Figura 2 ele apresenta uma sequência esquemática desde o dispositivo até a aplicação em si, utilizada para avaliar o processo com a otimização das rotas dos caminhões de mineração e no controle da velocidade dos caminhões para maior segurança sem aumento no custo de mineração.

|  |
| --- |
| Uma imagem contendo Carta  Descrição gerada automaticamente |
| Figura 2: Device design and network  Fonte: H. Aguirre-Jofre et al. (2021) |

O estudo aqui apresentado, também faz uso de dispositivos de IoTs de custo significativamente baixos, com uma camada de análise estática e algoritmos que buscam otimizar o processo de análise dos dados, usando de técnicas de Machine Learning na identificação de variações nos padrões de comportamento consideradas anômalos, demonstrando o potencial do CMMS e da integração de tecnologias avançadas para otimizar o monitoramento de condições em componentes críticos de equipamentos de mineração.

J.P. Dias et al. (2022) em seu artigo “Projetando e construindo sistemas de internet das coisas: um Visão geral do ecossistema “, contribui com uma visão ampla e geral do atual estado da arte sobre como projetar e construir sistemas de IoT, e aponta com clareza os desafios de dessa área.

Na figura 3 e mostrado o aumento do número de dispositivos de IoT por ano, onde J.P. Dias et al. (2022) avaliam tanto o aumento considerável de projetos de pesquisa quando a quantidade de dispositivos, demonstrando o interesse tanto da academia quando das empresas/pessoas por esta solução.

|  |
| --- |
| Gráfico, Gráfico de linhas  Descrição gerada automaticamente |
| Figura 3: Contagem do número de dispositivos IoT por ano  Fonte: J.P. Dias et al. (2022) |

J.P. Dias et al. (2022) relembram que o termo Internet das Coisas foi cunhado por Kevin Ashton por volta de 1999 em uma apresentação sobre o gerenciamento da cadeia de suprimentos, e que posteriormente o termo foi apresentado pelo comitê técnico conjunto da International Organization for Standardization (ISO) e a Comissão Eletrotécnica Internacional (IEC), definiram Internet das Coisas como sendo uma infraestrutura composta por objetos, pessoas, sistemas e recursos de informação todos interconectados, acrescenta-se a isso os serviços inteligentes para permitir que eles processem informações sobre o mundo físico e virtual.

A instalação e configuração dos sensores no equipamento foi realizada, considerando a cobertura de 100% dos pequenos componentes, que não são cobertos pelo monitoramento do fabricante, tal como descrito posteriormente nas Figuras 3, a seguir.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Uma imagem contendo no interior, mesa, par, deitado  Descrição gerada automaticamente | | Uma imagem contendo mesa, fogo, sujo, velho  Descrição gerada automaticamente | | Uma imagem contendo hidrante, sujo, escuro, mesa  Descrição gerada automaticamente | | Garrafa de vidro  Descrição gerada automaticamente com confiança média | |
| Uma imagem contendo no interior, mesa, bolsa, deitado  Descrição gerada automaticamente | Uma imagem contendo no interior, mesa, cheio, em pé  Descrição gerada automaticamente | | Garrafa de vidro  Descrição gerada automaticamente com confiança média | | Uma imagem contendo fogo, escuro, hidrante, quarto  Descrição gerada automaticamente | |
| Figuras 4a, 4b, 4c, 4a, 4d, 4e, 4f, 4g, 4h, 4i. Instalação dos dispositivos IoT (Temperatura/Vibração) nos componentes em campo  Fonte: Próprio Autor | | | | | | | |

Estes são sensores que coleta e enviam para um repositório na nuvem, os dados de temperatura e vibração dos componentes, o detalhamento deste fluxo e a configuração estão mais bem descritos nos tópicos posteriores de 3.3. e 3.7, onde será descrito com detalhes a aplicação, coleta e tratamento dos dados.

Como informado anteriormente, este trabalho teve como estudo a utilização de análise de temperatura com câmera térmica, e que obteve resultados promissores em termos de monitoramento preditivo, e será explicado em detalhes a seguir neste estudo. Com base nesse trabalho anterior, avançamos com este estudo, a fim de expandir o método, empregando mecanismos de Internet das Coisas (IoT) para obter dados diretamente dos componentes, possibilitando a análise de dados para detectar alterações de temperatura e vibração antes de uma falha catastrófica.

**Armazenamento na Nuvem**

Neste estudo, considerando os desafios de implementação e o foco na implementação de soluções mais focado na camada de Data Science, foi definido pelo uso de um sensor que coleta temperatura e vibração e que já possui, todo o ecossistema de IoT desenvolvido, eliminando assim a necessidade de estabelece e implementar estes pontos, e permitindo avançar diretamente para a parte de Análise estatística dos Dados após estarem disponibilizados na Nuvem. Ou seja, existe boas opções de soluções compatíveis e disponíveis no mercado para coletar e disponibilizar os dados, sendo algumas de custo mais elevado, outras com mais facilidade de instalação e menor complexidade do possibilitando que o foco desse trabalho fosse direcionado para a análise e projeção dos dados após estarem disponíveis. Na Figura 5 tem uma estrutura básica de camadas de IoTs desde a coleta dos dados, armazenamento e posterior análise dos dados utilizando neste caso, a nuvem.

Como apontado também por J.P. Dias et al. (2022), existem pontos críticos nesta etapa do processo, ligados a heterogeneidade, distribuição lógica e geográfica, preocupações humanas, necessidades de comunicação em tempo real e restrições de energia desempenham um papel fundamental no projeto, desenvolvimento, testes e manutenção da IoT.

|  |
| --- |
| Diagrama  Descrição gerada automaticamente |
| Figura 5. Visão logica de uma Camada comum de Sistema IoT  Fonte: J.P. Dias et al. (2022) |

Com isso, definimos pela solução integrada da empresa nacional Dynamox, adquirimos os sensores para instalação em campo + o Gateway, possibilitando o foco de nosso estudo na camada de ETL dos dados.

O Sistema DynaPredict, coleta e disponibiliza em tempo real todos os dados via API da base da empresa no Google Could para nossa base na nuvem, também no Google Cloud, conforme apresentado na Figura 4. Posteriormente foi desenvolvida todo tratamento e análise dos dados de Temperatura e Vibração conforme será detalhado a seguir neste estudo, tendo como principal foco, as Bombas Principais da escavadeira Caterpillar 395.

A estrutura da Figura 6, já está predisposta para receber os dados de campo de cada um dos sensores de forma online, via o Gateway utilizando sistema GSN 2, 3 ou 4g.

|  |
| --- |
| **Interface gráfica do usuário, Tabela  Descrição gerada automaticamente** |
| Figura 6. Visão Arvore de componentes Nuvem de Dados  Fonte: Próprio Autor |

**Series temporais**

Aliene Nielsen (2021) descreve uma história sobre a importância das series temporais em nosso dia a dia, e correlaciona isso a evolução das IoTs, salientando a relevância e importância deles em nosso cotidiano, ele ressalta aumento exponencial do uso das series temporais inicialmente na medicina como tendo o pioneirismo na geração de dados, meteorologia e crescimento econômico vindo posteriormente bem como o mercado de trading e astronomia.

Anderson et al. (2023) avalia uma série temporal como sendo um conjunto de variáveis ordenadas em função do tempo, e que o estudo e modelagem destes dados, é fundamental para compreender o comportamento e realizar previsões sobre determinados dados. Ele orienta para a necessidade de realização de um estudo do tipo de serie temporal que está sendo objeto de análise e projeção, e que em caso de observar que os dados possuem tendência estocástica, é preciso avaliar pois isso pode tornar as previsões imprecisas conforme se afastam do último ponto da amostra. Para tanto, ele orienta a utilização do procedimento para a avaliação da estacionariedade da série, por meio do teste de Dickey-Fuller, como sendo um dos testes de raiz unitária mais tradicionais. A seguir, temos na Eq (1, 2 e 3) representando o teste de Dickey-Fuller.

|  |  |
| --- | --- |
|  |  |
|  | (1)  (2)  (3) |

em que 𝜙 é um parâmetro a ser estimado, 𝑢𝑡 é um processo de ruído branco, que se caracteriza como uma sequência de variáveis aleatórias independentes e identicamente distribuídas (iid), com média zero e variância constante (𝑢𝑡∼𝑅𝐵 (0, 𝜎2)), 𝛽1é uma constante que representa o intercepto e 𝛽2 é um efeito de tendência. O teste considera como hipótese nula, 𝐻0 ∶ 𝜙= 1, a presença de raiz unitária e como hipótese alternativa, 𝐻1∶𝜙 <|1|, a série sendo estacionária.

Silveira et al. (2022) Na presença de estacionariedade, a função amostral que descreve o processo gerador dos da dos tem a mesma forma em todos os instantes, facilitando a identificação de estimativas dos parâmetros desconhecidos dos modelos especificados, e relaciona alguns avanços nos testes de ADF e sugere a utilização do cálculo do valor- p como determinado a partir da estatística τ no teste ADF, e possíveis interferências no resultado.

Em nosso estudo, definiremos o valor-p (p-value) e em caso de valores próximo de zero, indica que há forte evidência estatística contra a hipótese nula de que a série temporal é não estacionária, ou seja, é possível concluir com confiança que a série temporal é estacionária.

Aileen Nielsen (2021) explica, que se a série temporal é estacionária simplifica significativamente a análise e modelagem pois indica que suas propriedades estatísticas, como média, variância e autocorrelação, permanecem constantes ao longo do tempo. Isso permite aplicar uma variedade de técnicas de modelagem e previsão com mais confiança.

Aliene Nielsen (2021) separa os modelos estatísticos de series temporais entre, estáticos, modelos de esparo de estado e aprendizado de máquina, sendo que em nosso estudo, utilizaremos alguns destes descritos a seguir, sendo que a avaliação e definição do modelo mais aderente estará condicionada aos resultados de MAE (), MSE e R2 para os seguintes modelos, conforme descrito por Aurélien Géron (2021) como sendo alguns dos métodos estatísticos desenvolvidos para series temporais:

Regressão Linear Simples ou Múltipla: Aurélien Géron (2021) A regressão linear é um modelo simples que pode ser eficaz se houver uma relação linear direta entre as variáveis de entrada e saída.

Modelos Autorregressivos (AR): Modelos autorregressivos consideram a relação entre uma observação atual e observações passadas. Eles são úteis quando há dependência temporal nas séries temporais.

Médias Móveis (MA): Modelos de médias móveis consideram a relação entre uma observação e um erro residual das observações passadas. Eles podem ser combinados com modelos AR para formar modelos ARMA.

Modelos ARIMA (Autoregressive Integrated Moving Average): Aurélien Géron (2021) avalia que os modelos ARIMA combinam componentes de regressão autorregressiva, médias móveis e diferenciação para modelar séries temporais estacionárias ou com tendências conhecidas.

Suavização Exponencial (Exponential Smoothing): Modelos recomendados para suavizar séries temporais e capturar padrões sazonais.

Redes Neurais Recorrentes (RNNs): Como dito por Aurélien Géron (2021), modelos de redes neurais recorrentes, como LSTM e GRU, são eficazes para capturar dependências de longo prazo em séries temporais, sendo a LSTM tem resultados melhores, contudo a GRU e mais rápida no treinamento mesmo que em determinados estudos o resultado da GRU tenha sido melhor.

State Space Models: Modelos recomendados para modelar séries temporais que podem ser decompostas em estados latentes.

Também como concluído por Aliene Nielsen (2021), modelos complicados nem sempre são os melhores, pois o custo-benefício compensa o uso de recursos computacionais adicionais requeridos para sua operação.

Em uma análise e definição de um modelo, e importante pensar se o tempo de treinamento adicional para operar com um modelo de aprendizagem de máquinas complexo valem a pena.

Aliene Nielsen (2021) também apresenta as diferenças entre os modelos de series temporais univariadas e multivariadas, ou seja, series temporais univariadas possuem apenas uma variável medida ao longo do tempo, já as series temporais multivariadas são series com múltiplas variáveis medidas a cada timestamp. Esta última, são muito uteis para análise, pois muitas vezes as variáveis calculadas são inter-relacionadas, e mostram dependências temporais entre si.

Neste estudo, a serie temporal base para o trabalho de Data Science, possui dados de temperatura e vibração de pequenos componentes da escavadeira Caterpillar 395, contudo, a análise se concentrará em somente uma variável ao longo do tempo.

**Trabalhos relacionados/Estudo de caso**

Inicialmente, e como comentado acima, este estudo foi aplicado sem quaisquer usos de tecnologia embarcada, e com a utilização de um inspetor e uma câmera tecnográfica.

Este procedimento operacional requerido anteriormente, demandava uma série de atividades obrigatórias para a compreensão de algum desvio nos parâmetros, sendo seguidas todas as etapas de avaliação de desempenho seguindo as diretrizes do fabricante tais como:

* Tempo de descida e subida do cilindro,
* Pressão hidráulica durante os testes,
* Taxa de fluxo das bombas hidráulicas.

As Figuras 7a e b representam as pressões hidráulicas do equipamento, antes de realizar o teste de velocidade de descida e subida do implemento.

|  |  |
| --- | --- |
| Interface gráfica do usuário  Descrição gerada automaticamente | Interface gráfica do usuário, Aplicativo  Descrição gerada automaticamente |
| Figura 7a. Temperatura do Fluido e Figura 7b. Pressão Hidráulica  Fonte: Próprio autor | |

Somente após estes pontos estarem dentro do parâmetro especificação, o teste de temperatura por termografia foi conduzido, em paralelo aos testes de tempo de ciclo e pressão.

Utilizando uma câmera termográfica para registrar o diferencial de temperatura dos cilindros eram feitas medições nas temperaturas, procurando identificar possíveis diferencial térmico foi entre os cilindros, neste caso, e conforme ilustrado na Figura 8, foi possível identificar um diferencial de temperatura de 4,7°C em relação ao lado esquerdo.

|  |
| --- |
| Uma imagem contendo no interior, mesa, diferente, colorido  Descrição gerada automaticamente |
| Figura 8. Termografia nos Cilindros de Elevação da Escavadeira  Fonte: Próprio autor |

Os pontos mais quentes dos cilindros foram capturados pela câmera, revelando uma diferença de temperatura de ~5°C entre o cilindro de elevação do lado direito (l/d) e o cilindro de elevação do lado esquerdo (l/e) (Figura 9).

|  |
| --- |
| Imagem de vídeo game  Descrição gerada automaticamente com confiança média |
| Figura 9. Termográfica com 5°C entre os pontos dos cilindros de elevação da lança  Fonte: Próprio autor |

Após remover e desmontar o cilindro na oficina de reparos, foi possível confirmar que os sintomas observados no campo foram importantes para definir e concluir que o método termográfico pode ser usado com mais segurança para determinar a necessidade de remoção, conforme mostra a Figura 10 a seguir. Isso resultou em uma redução, embora não significativa que iremos explicar mais a seguir, nos custos de reparo e no impacto na contaminação do sistema hidráulico.

|  |
| --- |
| Uma imagem contendo trem, caminhão, mesa, motor  Descrição gerada automaticamente |
| Figura 10. Cilindro na Centro de Reforma  Fonte: Próprio autor |

As Figuras 11a, 11b e 11c fornecem detalhes adicionais após a desmontagem e análise do cilindro, sendo que o modo de falha apresentado na Imagem 7 ilustra a causa do aumento de temperatura no processo termográfico de campo.

|  |  |  |
| --- | --- | --- |
| Uma imagem contendo no interior, armário, cozinha, aço  Descrição gerada automaticamente | Uma imagem contendo ao ar livre, edifício, metal, relógio  Descrição gerada automaticamente | Imagem em preto e branco em cima de uma mesa de madeira  Descrição gerada automaticamente com confiança baixa |
| Figura 11a, b e c. Falha interna cilindro após desmontagem e peritagem  Fonte: Próprio autor | | |

Com isso, o que foi possível entender é que mesmo tendo sido avaliado antecipadamente, a frequência em que foram realizadas as inspeções e coletas em campo permitiram um avanço da falha até um nível que o sistema já havia sido contaminado, apesar de ter sido removido antes de uma falha catastrófica, tal abordagem permitiu avanço nos significativos na parte interna do componente.

Após os resultados utilizando fluxo anterior, e consequentemente com base na conclusão após desmontagem e reforma, o principal questionamento a ser respondido foi, qual seria a forma de avançar neste processo de coleta e análise dos dados de temperatura a tempo de remover o componente, sem que ele possa estar em um estado de degradação avançado? Como mitigar isso? Como antecipar a falha, a ponto de não deixar esse Modo de Falha ocorrer?

**Fluxograma de análise dos dados**

As etapas deste trabalho são descritas a seguir no Figura 12, desde a análise inicial até a conclusão, com base em uma Estrutura Analítica do Projeto (EAP) simples e didatica.

Figura 12. EAP projeto produto/processo

Fonte: Próprio autor

**Definição dos requisitos do projeto Brainstorm**

O projeto anterior, apresentado como parte de uma estratégia de evoluir no monitoramento de pequenos componentes em campo, alguns pontos de evolução identificados, e são apresentados a seguir na Figura 13 como sendo as possíveis causas do avanço na falha potencial do cilindro anteriormente analisado, e sugerido ações para mitigar tais anormalidades.

|  |
| --- |
| Interface gráfica do usuário, Texto, Aplicativo, Email  Descrição gerada automaticamente |
| Figura 13. Ishikawa: Análise possíveis causas  Fonte: Próprio autor |

**Definição da tecnologia IoT**

Consequentemente, uma proposta de solução foi requerida, e esquematizada na Figura 14, com sendo um aprimoramento no processo de monitoramento, a fim de estruturar uma análise robusta, com uso de tecnologia de sensoriamento online e métodos estatísticos e suporte decisório mais sólido.

|  |
| --- |
| Diagrama  Descrição gerada automaticamente |
| Figura 14. Online Data Collect  Fonte: Próprio autor |

Já o equipamento/modelo definido para o piloto deste estudo em campo, foi a Escavadeira Hidráulica Caterpillar 395 primeiramente, por ter sido o modelo de ativo utilizado no primeiro trabalho, utilizando a câmera termográfica apesentado na introdução desse estudo, segundo pela quantidade de ativos deste modelo em campo e terceiro pela sua criticidade para a produção nos clientes.

Já os componentes, foram todos os pequenos componentes que não possuem monitoramento direto pelo sistema embarcado do ativo.

Na Figura 15 temos a escavadeira Caterpillar 395 e os componentes definidos para monitoramento online.

|  |
| --- |
| Grupo de música no palco com instrumentos musicais e microfones  Descrição gerada automaticamente com confiança baixa |
| Figura 15. Escavadeira 395 e componentes a serem monitorados  Fonte: Próprio autor |

Já para gerenciamento dos parâmetros, foi definido uma estrutura de Sistema/Componentes conforme apresentado na Figura 16.

A Figura 15 a seguir, temos a estrutura hierárquica dos componentes a ser monitorados.

|  |
| --- |
| Texto  Descrição gerada automaticamente |
| Figura 16. Estrutura de sistemas/Componentes Modelo 395  Fonte: Próprio autor  \*\* Em verde, o componente definido para análise neste estudo |

**Pontos de Instalação e Coleta**

Para a análise e apresentação dos resultados, foi utilizando alguns modelos de análise estatística de series temporais, com foco nos dados coletados para os componentes Bomba Principal P1 e P2, como apresenta a Figura 17 a seguir, tanto por ser o componente mais crítico para o funcionamento do ativo pelo custo, impacto na disponibilidade física e por causar um dado consequentes alto quando em uma parada não programada.

Os dados de temperatura e vibração da bomba principal foram separados e tratados conforme apresentado a seguir, para posteriormente, ser elaborada a abordagem de Machine Learning mais apropriada.

|  |
| --- |
| Diagrama  Descrição gerada automaticamente com confiança média |
| Figura 17. Componentes monitorados  Fonte: Próprio autor |

**ETL dados de temperatura Bombas Principais 1 e 2.**

A escavadeira de mineração Caterpillar 395 foi selecionada por ser um ativo importante para a produção do cliente, oferecendo versatilidade e por ter sido o mesmo ativo utilizado no projeto anterior, onde a coleta foi feita manualmente.

Assim, foram instalados os sensores nos componentes, conforme apresentado na Figura 18, e feita a integração com a base de dados da Dynamox, posteriormente foi feita a integração dos dados disponibilizadas na Cloud que iremos utilizar para a análise dos dados de temperatura apresenta o comportamento apresentado no Figura 18 a seguir:

|  |
| --- |
| Linha do tempo  Descrição gerada automaticamente |
| Uma imagem contendo Gráfico  Descrição gerada automaticamente |
| Figura 18. Distribuição das temperaturas das Bombas 1 e Bombas 2  Fonte: Próprio Autor |

A base de dados original fornecida pela plataforma da Dynamox possui a estrutura mostrada na Figura 19 a seguir, sendo necessário um tratamento nos dados para seguirmos com a análise e projeção:

|  |
| --- |
| Parametro Timestamp Value NmeComp  263928 AccelerationX 11-Apr-2022 23:53:52 0.2 MainPumpP1  263929 AccelerationZ 11-Apr-2022 23:53:52 0.2 MainPumpP1  263930 AccelerationY 11-Apr-2022 23:53:52 0.2 MainPumpP1  ... ... ... ... ...  394843 Temperature 28-Apr-2022 15:21:12 84.2 MainPumpP2  394844 Temperature 28-Apr-2022 15:24:12 85.2 MainPumpP2  394845 Temperature 28-Apr-2022 15:27:12 85.2 MainPumpP2 |
| Figura 19. Estrutura dos dados de temperatura e vibração  Fonte: Próprio Autor |

Já na Figura 20 foram removidos os dados de Vibração para ambas as Bombas, para que este estudo de análise de dados possamos focar somente na temperatura de ambas as bombas.

|  |
| --- |
| Parametro Timestamp Value NmeComp  287912 Temperature 11-Apr-2022 12:25:10 28.5 MainPumpP1  287913 Temperature 11-Apr-2022 12:28:10 28.4 MainPumpP1  287914 Temperature 11-Apr-2022 12:31:10 28.3 MainPumpP1  ... ... ... ... ...  394845 Temperature 28-Apr-2022 15:27:12 85.2 MainPumpP2  394846 Temperature 28-Apr-2022 15:30:12 85.3 MainPumpP2  394847 Temperature 28-Apr-2022 15:33:12 85.4 MainPumpP2 |
| Figura 20. Estrutura dos dados somente com a temperatura  Fonte: Próprio Autor |

Para estratificação e análise dos dados foi utilizado o Software Python com interface do Google Colab. Iniciando pela avaliação dos dados gerais, apresentado nas Figuras 21 e 22a seguir para a temperatura da Bomba Principal 1 e 2.

Os dados analisados e apresentados no Figuras 21 a seguir, são de um período de 17 dias de coleta de temperaturas, iniciando em 11/03/2023 até 28/04/2023, compondo 8226 observações.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **NmeComp** | **Value count** | **mean** | **std** | **min** | **25%** | **50%** | **75%** | **max** | | **MainPumpP1** | 8226 | 52,2 | 16,6 | 24,2 | 37,6 | 57,0 | 62,0 | 99,9 | | **MainPumpP2** | 8226 | 76,7 | 17,3 | 38,4 | 65,6 | 81,0 | 90,7 | 104,0 | |
| Figuras 21. Resumo dados temperatura Bomba 1 e 2  Fonte: Próprio autor |

Já a dispersão dos dados é apresentada no Figuras 22 e 23 a seguir, são de um período separadas pela bomba 1 e bomba 2,

|  |
| --- |
| Gráfico, Histograma  Descrição gerada automaticamente |
| Figuras 22a e b. Perfil temperatura Bomba principal 1 e 2  Fonte: Próprio autor |

Na amostra coletada para este estudo, observou-se uma concentração de valores de temperatura formando uma assíntota a esquerda para os dados da Bomba 1, e uma assíntota a direita para os dados da Bomba 2.

Avançando na análise dos dados, utilizamos o pacote “scikit-learn” que como apresentado no livro Géron, A. (2019) são capazes de realizar tarefas de análise de dados de regressão.

**Modelagem matemática**

A seguir as Figuras 23a e b, mostra o teste de Dickey-Fuller, utilizando a biblioteca do python “*import adfuller*” para identificar o padrão estacionário ou não dos dados. Os resultados dessa análise para a amostra de dados do estudo, apresentaram um Valor-p = 0.000003 para a Bomba P1 e de Valor-p = 0.000009 para a Bomba P2, estão muito próximo de zero, ou seja, indicando que há forte evidência estatística para rejeitar a hipótese nula de que a série é não estacionária, ou seja, este resultado tanto para a Base de dados da Bomba 1 quanto para a Bomba 2, sugere-se que a série temporal seja estacionária.

|  |  |
| --- | --- |
| Interface gráfica do usuário, Gráfico  Descrição gerada automaticamente | Gráfico  Descrição gerada automaticamente |
| Figuras 23a. Valor-p para dados Bomba 1  Fonte: Próprio Autor | Figuras 23b. Valor-p para dados Bomba 2  Fonte: Próprio Autor |

A definição pela técnica de AST&P (Advanced Statistical Techniques & Procedures) para análise de dados e tomada de decisão foram feitos utilizando técnicas estatísticas clássicas e avançadas. Dentre estes, o modelo mais aderente para trabalhar com séries temporais lineares, depende de vários fatores, incluindo a estrutura dos dados, a quantidade de dados disponíveis, a presença de tendências ou sazonalidades, entre outros, contudo os modelos que foram testados para avaliar os resultados nessa série temporal de temperatura foram:

* *Regressão Linear Simples ou Múltipla*: A regressão linear é um modelo simples que pode ser eficaz se houver uma relação linear direta entre as variáveis de entrada e saída.
* *Modelos Autoregressivos (AR)*: Modelos autorregressivos consideram a relação entre uma observação atual e observações passadas. Eles são úteis quando há dependência temporal nas séries temporais.
* *Médias Móveis (MA):* Modelos de médias móveis consideram a relação entre uma observação e um erro residual das observações passadas. Eles podem ser combinados com modelos AR para formar modelos ARMA.
* *Modelos ARIMA (Autoregressive Integrated Moving Average):* Modelos ARIMA combinam componentes de regressão autorregressiva, médias móveis e diferenciação para modelar séries temporais estacionárias ou com tendências conhecidas.
* *Suavização Exponencial (Exponential Smoothing):* Modelos recomendados para suavizar séries temporais e capturar padrões sazonais.
* *Redes Neurais Recorrentes (RNNs):* Modelos de redes neurais recorrentes, como LSTM e GRU, são eficazes para capturar dependências de longo prazo em séries temporais.
* *State Space Models:* Modelos recomendados para modelar séries temporais que podem ser decompostas em estados latentes.

Com base nos resultados obtidos tanto no teste de Dickey-Fuller quanto nos resultados de MAE, MSE e R2, como apresentado na Figuras 24, com os resultados consolidados dos modelos, e possível identificar quais os modelos apresentam maior aderência aos dados coletados.

|  |  |  |  |
| --- | --- | --- | --- |
| Modelo | MAE | MSE | R2 |
| State Space Model | **1.294.020.169** | **2,25E+12** | **-4.729.475.731** |
| SARIMAX | **13.243.707** | **2,74E+08** | **-0.012671** |
| RNN-GRU | **0.028991** | **2,43E+03** | **0.948623** |
| RNN-LSTM | **0.026283** | **2,26E+03** | **0.952317** |
| Regressão Linear Múltipla | **12.297.906** | **2,15E+08** | **0.206178** |
| Média Movel | **2.440.817** | **1,50E+07** | **0.945099** |
| Model AR | **5.576.465** | **8,25E+07** | **0.826393** |
| ARIMA | **22.282.574** | **6,97E+08** | **-0.466564** |
| SVR | **9.104.204** | **1,49E+08** | **0.450116** |
| SVR\_Ajustado | **1.866.355** | **1,36E+07** | **0.943082** |
| Exponential Smoothing | **16.322.945** | **4,44E+08** | **-0.643456** |
| Correlação | **1.827.380** | **1,81E+07** | **0.933978** |

|  |
| --- |
| Figuras 24. Resumo dados temperatura Bomba 1  Fonte: Próprio Autor |

As métricas de desempenho para os diferentes modelos de previsão de séries temporais aplicados aos dados de temperatura das bombas P1, foram avaliados em termos de três métricas:

* **MAE (Mean Absolute Error):** Erro absoluto médio, que mede a média dos erros absolutos entre os valores previstos e os valores reais. Valores menores indicam melhor desempenho.
* **MSE (Mean Squared Error):** Erro quadrático médio, que mede a média dos quadrados dos erros entre os valores previstos e os valores reais. Valores menores indicam melhor desempenho.
* **R² (Coeficiente de Determinação):** Mede a proporção da variação dos dados que é explicada pelo modelo. Valores próximos de 1 indicam um modelo que explica bem a variação dos dados, enquanto valores negativos indicam um modelo que está performando pior do que uma simples média.

State Space Model: MAE: 1294.020169, MSE: 2.247582e+06, R²: -4729.475731, os resultados mostram que este modelo tem um desempenho extremamente ruim. O MAE é muito alto, indicando previsões imprecisas. O MSE também é extremamente alto, e o R² negativo indica que o modelo está performando muito pior do que a média.

SARIMAX: MAE: 13.243707, MSE: 2.736540e+02, R²: -0.012671, os resultados mostram que o modelo SARIMAX também não performa bem, com um R² ligeiramente negativo e um MAE e MSE altos, indicando que não é um bom modelo para esses dados.

RNN-GRU: MAE: 0.028991, MSE: 2.431490e-03, R²: 0.948623 os resultados mostram que o modelo RNN-GRU apresenta um excelente desempenho, com um MAE muito baixo, um MSE extremamente baixos e um R² muito próximo de 1, indicando que o modelo explica bem a variação dos dados.

RNN-LSTM: MAE: 0.026283, MSE: 2.256681e-03, R²: 0.952317, os resultados mostram que o modelo RNN-LSTM também apresenta um desempenho excelente, similar ao GRU, com um MAE, MSE e R² muito bons.

Regressão Linear Múltipla: MAE: 12.297906, MSE: 2.145145e+02, R²: 0.206178 os resultados mostram que este modelo tem um desempenho mediano, com um MAE e MSE relativamente altos e um R² baixo, indicando que não explica bem a variação dos dados.

Média Móvel: MAE: 2.440817, MSE: 1.502500e+01, R²: 0.945099, os resultados mostram que a técnica de média móvel tem um bom desempenho, com um MAE razoavelmente baixo, um MSE baixo e um R² alto.

Modelo AR: MAE: 5.576465, MSE: 8.251220e+01, R²: 0.826393, os resultados mostram que o modelo AR (Autorregressivo) tem um desempenho decente, com um MAE e MSE moderados e um R² relativamente alto.

ARIMA: MAE: 22.282574, MSE: 6.968059e+02, R²: -0.466564, os resultados mostram que o modelo ARIMA não performa bem, com um MAE e MSE altos e um R² negativo, indicando desempenho ruim.

SVR: MAE: 9.104204, MSE: 1.485951e+02, R²: 0.450116, os resultados mostram que o modelo SVR tem um desempenho razoável, com um MAE e MSE moderados e um R² relativamente baixo.

SVR\_Ajustado: MAE: 1.866355, MSE: 1.356212e+01, R²: 0.943082, os resultados mostram que o modelo SVR\_1 apresenta um desempenho muito bom, com um MAE baixo, um MSE baixo e um R² alto.

Exponential Smoothing: MAE: 16.322945, MSE: 4.441108e+02, R²: -0.643456, os resultados mostram que o modelo de suavização exponencial não performa bem, com um MAE e MSE altos e um R² negativo.

Correlação: MAE: 1.827380, MSE: 1.808035e+01, R²: 0.933978, os resultados mostram que a técnica de correlação apresenta um bom desempenho, com um MAE baixo, um MSE relativamente baixo e um R² alto.

Neste caso, os modelos de RNN-GRU e RNN-LSTM obtiveram resultados bem mais aderentes ao modelo, com os menores MAE e MSE e os R² mais altos. Alinhado a isso, também o modelo SVR\_Ajustado com R*andomized Search CV*, como técnica de otimização de hiperparâmetros do pacote scikit-learn, para encontrar a combinação ideal de hiperparâmetros para um modelo de machine learning, maximizou a performance e apresentou bom desempenho a ponto de ajusta este modelo como sendo um dos mais aderentes para projeção dos dados.

Esta projeção tem como objetivo fundamental, a identificação de quando os valores de temperatura podem atingir o limite de temperatura que indique uma falha potencial, conforme pode ser visto na Figuras 25 e Figuras 26.

**Resultados e Discussão**

A projeção com o modelo LSTM foi elaborada utilizando os resultados do MAE, MSE e R-Quadrado apresentados na Figura 24. O modelo empregou a função de ativação tanh, com regularizadores, kernel\_regularizer L1(0.01), recurrent\_regularizer L2(0.01) e bias\_regularizer L2(0.01). Além disso, um Dropout de 0.2 foi aplicado para prevenir overfitting, e o otimizador utilizado foi o "adam".

Para a análise e pré-processamento dos dados, utilizamos o pacote scikit-learn. A projeção foi realizada para os 25 dias subsequentes, baseando-se em um conjunto de 28 dias de dados coletados. Com base nos dados históricos, foi traçada uma linha representando o limite superior da temperatura, servindo como uma referência para as projeções futuras.

Os resultados da projeção para a LSTM não foram satisfatórios com base no objetivo proposto, dado que o resultado da projeção demostra uma convergência para valores após 25 dias futuros, o que impediu a utilização deste modelo para atingir o objetivo de entender, quando os valores de temperatura atingiriam o limite superior, com base nos dados históricos, como pode ser observado na Figura 2 a seguir.

|  |
| --- |
| **Gráfico  Descrição gerada automaticamente** |
| Figuras 25. Projeção de temperatura Bomba 1  Fonte: Próprio Autor |

Para o modelo GRU, os resultados da projeção foram basicamente simulares, tendo como projeção uma convergência para valores após 25 dias futuros, o que impediu a utilização deste modelo para atingir o objetivo de entender, quando os valores de temperatura atingiriam o limite superior, com base nos dados históricos, como pode ser observado na Figura 26 a seguir.

|  |
| --- |
| **Gráfico  Descrição gerada automaticamente** |
| Figuras 26. Projeção de temperatura Bomba 2  Fonte: Próprio Autor |

Já para o modelo de SVM, com RandomizedSearchCV de otimização de hiperparâmetros para ajuste do modelo, obtendo resultados demonstrados na tabela 24 de ajuste.

Com os ajustes, o modelo Suport Vector Regression, foi possível identificar que na projeção para os próximos 25 dias, os valores de temperatura cruzariam o limite superior de temperatura na data de 03/março, mais precisamente as 15:33 minutos.

|  |
| --- |
| **Gráfico  Descrição gerada automaticamente** |
| Figuras 27. Projeção de temperatura Bomba 2  Fonte: Próprio Autor |

Avançando na análise dos resultados, realizamos o teste de Shapiro-Wilk para GRU, LSTM e SVR com o objetivo de compreender melhor os resíduos dos modelos e o que eles indicam sobre os resultados.

|  |  |
| --- | --- |
| **Parametro** | **Análise** |
| Média dos Resíduos | **SVR:** A média dos resíduos é de -0.10, próxima de zero, indicando que as previsões, em média, não estão significativamente deslocadas dos valores reais.  **LSTM:** A média dos resíduos é de 0.31, também próxima de zero, sugerindo uma boa precisão nas previsões.  **GRU:** A média dos resíduos é de 1.08, ligeiramente maior, indicando um pequeno viés nas previsões. |
| Dispersão dos Resíduos | **LSTM:** O desvio padrão dos resíduos é de 3.75, sugerindo que as previsões do LSTM são mais consistentes e têm menor variabilidade.  **SVR:** O desvio padrão é de 6.46, indicando maior variabilidade nas previsões, sugerindo menor consistência.  **GRU:** O desvio padrão é intermediário, entre LSTM e SVM, mostrando uma variabilidade moderada nas previsões |
| Resíduos Extremos | **SVR:** Apresenta os resíduos mínimos e máximos mais extremos, sugerindo a presença de outliers ou erros de previsão mais significativos.  **GRU e LSTM:** Têm menores valores de resíduos extremos em comparação ao SVR, com o LSTM apresentando os menores valores de resíduos extremos, indicando previsões mais controladas e menos outliers |
| Quartis | **SVR:** A maior dispersão interquartil (IQR), indicando uma maior variabilidade entre o primeiro e o terceiro quartil dos resíduos.  **GRU:** Apresenta uma dispersão interquartil intermediária.  **LSTM:** Tem a menor dispersão interquartil, sugerindo uma distribuição mais compacta dos resíduos e, portanto, previsões mais consistentes |

Ou seja, os resultados da análise indicam que o modelo LSTM apresenta previsões mais consistentes e menos variáveis, com resíduos próximos de zero e menor dispersão. O SVR, apesar de ter uma média de resíduos próxima de zero, apresenta maior variabilidade e resíduos extremos mais significativos, indicando a presença de outliers. O GRU mostra um desempenho intermediário entre o LSTM e o SVR.

|  |
| --- |
| **Gráfico, Histograma  Descrição gerada automaticamente** |
| **Gráfico, Gráfico de dispersão  Descrição gerada automaticamente** |
| Figuras 28. Teste dos resíduos para os modelos GRU, LSTM e SVR  Fonte: Próprio Autor |

Os resultados obtidos para a amostra de dados coletados demonstraram diferentes comportamentos entre os modelos estudados. Os modelos de aprendizado de máquina, especificamente RNN-GRU e RNN-LSTM, destacaram-se por seu excelente desempenho em termos de métricas de erro (MAE e MSE) e explicação da variância (R2). Esses modelos conseguiram capturar padrões complexos nos dados temporais, tornando-os altamente eficazes.

Um modelo mais simples, a Média Móvel, também apresentou bom desempenho, com baixas métricas de erro e um R2 alto. Isso sugere que, em alguns casos, a média móvel pode ser uma abordagem eficaz para a previsão de séries temporais.

Em contrapartida, o modelo ARIMA teve um desempenho inferior, com altos valores de MAE e MSE e um R2 negativo. Isso indica que o ARIMA não conseguiu capturar adequadamente a estrutura dos dados de temperatura da Bomba 1.

Os modelos de Regressão Linear Múltipla e SVR tiveram um desempenho intermediário, apresentando valores moderados de MAE, MSE e R2. Esses modelos podem necessitar de ajustes adicionais para melhorar sua adequação aos dados.

Os modelos State Space Model, SARIMAX e Exponential Smoothing apresentaram desempenho inferior, com altos valores de MAE e MSE e R2 negativos. Esses modelos mostraram-se inadequados para capturar a complexidade dos dados temporais de temperatura da Bomba 1.

Considerando o conjunto de dados deste estudo, os modelos RNN-GRU, RNN-LSTM e SVR com Randomized Search CV foram os mais adequados, apresentando os melhores ajustes de MAE, MSE e R2. No entanto, é importante notar que os tempos de processamento dos modelos RNN-GRU e RNN-LSTM foram consideravelmente altos, indicando a necessidade de avaliar o impacto desse tempo na análise contínua dos dados

**Conclusões**

A análise dos dados de temperatura das bombas P1 e P2, com a aplicação de diversos modelos de previsão de séries temporais, resultou em insights valiosos sobre o desempenho dos modelos testados. Os modelos RNN-GRU e RNN-LSTM demonstraram um desempenho superior, evidenciado pelos menores valores de MAE e MSE e altos valores de R², indicando uma forte capacidade de captura dos padrões complexos nos dados. Esses modelos são recomendados para aplicações que requerem alta precisão na previsão de séries temporais, apesar do maior tempo de processamento envolvido.

A técnica de SVR com RandomizedSearchCV também apresentou resultados promissores, posicionando-se como uma alternativa robusta e eficiente, especialmente após a otimização de hiperparâmetros. Esse modelo mostrou-se adequado para projeções a curto prazo, como a identificação de cruzamento dos limites de temperatura que indicam possíveis falhas.

Por outro lado, métodos tradicionais como ARIMA e Suavização Exponencial não conseguiram captar a estrutura complexa dos dados de temperatura, resultando em maiores erros de previsão e menores valores de R². Isso sugere que esses modelos não são adequados para esse tipo específico de análise temporal.

A utilização de técnicas avançadas, como o RandomizedSearchCV para otimização de hiperparâmetros, foi essencial para melhorar o desempenho dos modelos, destacando a importância de um processo meticuloso e iterativo de ajuste de modelos em problemas de análise de dados industriais.

Em resumo, este estudo demonstrou que o uso de técnicas avançadas de aprendizado de máquina, combinadas com métodos estatísticos rigorosos, pode fornecer previsões precisas e úteis para a manutenção preditiva de equipamentos de mineração. A aplicação desses modelos permite uma melhor compreensão e antecipação das falhas, contribuindo para a eficiência operacional e redução de custos com manutenções emergenciais.

Para futuras pesquisas, recomenda-se a exploração de modelos híbridos que combinam diferentes abordagens para capturar a complexidade dos dados ainda mais eficientemente, bem como a aplicação desses métodos a um conjunto mais amplo de componentes e variáveis operacionais.

Esses resultados reforçam a viabilidade e importância da integração de tecnologias de IoT e machine learning no monitoramento e manutenção de ativos industriais, promovendo avanços na segurança de quem executa, confiabilidade e performance operacional.

**Referências**

Aguirre-Jofré, H. Eyre, M.Valerio, S. Vogt, D 2021. Low-cost internet of things (IoT) for monitoring and optimising mining small-scale trucks and surface mining shovels. Camborne School of Mines, University of Exeter, Exeter, United Kingdom and ndependent Consultant, Dataquest 131: 103918.

Aileen Nielsen 2021. Análise prática de séries temporais: predição com estatística e aprendizado de máquina. Alta Books: Rio de Janeiro

Artur Skoczylasa,Paweł Stefaniaka, Wiesława Gryncewiczb, Artur Rotb 2023. The Concept of an Intelligent Decision Support System for Ore Transportation in Underground Mine. 27th International Conference on Knowledge Based and Intelligent Information and Engineering Sytems 2023: 922–931.

Aurélien Géron 2021. Maos à Obra: Aprendizagem de Máquina com Scikit-Learn, Keras e TensrFlow - 2nd edição atualizada com TensorFlow2. O'Reilly Media, Inc.: Rio de Janeiro.

Daniel Schmidt, Karsten Berns 2013. Climbing robots for maintenance and inspections of vertical structures - A survey of design aspects and technologies. Robotics and Autonomous Systems 61: 1288-1305.

FÁVERO, Luiz Paulo Lopes e BELFIORE, Patrícia Prado 2024. Manual de Análise de Dados – Estatística e Machine Learning com Excel®, SPSS®, Stata®, R® e Python®. GEN LTC: Rio de janeiro

Ferreira, B., Seruffo, M., & Pires, Y. 2022. Planejamento e construção de um protótipo de aplicativo mobilem para visualização de dados de sistema de monitoramento de máquinas e equipamentos. Revista Principia - Divulgação Científica e Tecnológica do IFPB 59(3): 947-966.

Gbadamosi, Abdul-Quayyum & Oyedele, Lukumon & Davila Delgado, Manuel & Kusimo, Habeeb & Akanbi, Lukman & Olawale, Oladimeji & Muhammed -Yakubu, Naimah 2021. IoT for predictive assets monitoring and maintenance: An implementation strategy for the UK rail industry. Automation in Construction 122: 103486. 10.1016/j.autcon.2020.103486.

João Pedro Dias, André Restivo, Hugo Sereno Ferreira 2022. Designing and constructing internet-of-Things systems: An overview of the ecosystem. Internet of Things. journal homepage: www.elsevier.com/locate/io 19: 100529.

Manh-Kien Tran, Satyam Panchal, Vedang Chauhan, Niku Brahmbhatt, Anosh Mevawalla, Roydon Fraser, Michael Fowler 2021. Python-based scikit-learn machine learning models forthermal and electrical performance prediction of high-capacity lithium-ion battery. Int J Energy Res.2022 46: 786–794.

Nascimento, E. dos S., Maroli, K. R., Dias, G. C. M., & Bovério, M. A. 2020. GESTÃO DE MANUTENÇÃO INDUSTRIAL: eletrificação de acionamento de terno de moenda. SITEFA 3(1): 180–191.

Pedro A. Morettin, Clélia M. C. Toloi 2020. Análise de Series Temporais Modelos multivariados e não lineares. Edgard Blucher : São Paulo, SP.

scikit-learn.org scikit-learn. 2024 scikit-learn: Machine Learning in Python: <https://scikit-learn.org/stable/>. Acesso em 02 Jul. 2024

Yingming Tian, Fan Gao and Peng Wu2 1992. Intelligent Diagnosis of Equipment Health Based on IOT and Operation Large Data Analysis. Journal of Physics: Conference Series : 1742-6596.