**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 aborda 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, em que 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 aos fluxos manuais utilizados ainda com maior frequência 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 GRU (Gated Recurrent Unit), LSTM (Long Short-Term Memory), e SVR (Support Vector Regression), com a projeção de temperatura. Entre os principais resultados, observou-se que os modelos de aprendizado profundo, especialmente GRU e 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 GRU (Gated Recurrent Unit), LSTM (Long Short-Term Memory), e SVR (Support Vector Regression), projecting when components would reach critical operational limits. The main findings showed that deep learning models, particularly GRU and 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) escrevendo sobre 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 relacionados, 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.

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 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/tibração) 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, onde será descrito com detalhes a aplicação, coleta e tratamento dos dados.

Como informado anteriormente, este trabalho teve inspiração na análise de temperatura com câmera térmica, elaborado desde 2018, 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 camada de Data Science, para tanto, 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 Dynamox, adquirimos os sensores para instalação em campo + o Gateway, possibilitando o foco de nosso estudo na camada de ETL e Data Science e Analityc.

O Sistema DynaPredict, coleta e disponibiliza os dados via API da base da empresa no Google Could para nossa base na nuvem, também no Google Cloud, conforme apresentado na figura 6. 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 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, RMSE, AIC e R2 para os modelos com melhor performance, conforme descrito por Aurélien Géron (2021) como sendo alguns dos métodos estatísticos desenvolvidos para series temporais.

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.

Durante a avaliação dos dados, foi identificado que, ao se obter uma variação pequena na temperatura dos componentes, se obtinha um impacto significativo nos resultados das métricas avaliadas. Com base nisso, foi utilizado a transformação de Box Cox para trabalhar na estabilização da variância e tornar os dados mais simétricos, o que é particularmente útil quando a série temporal original exibe variância não constante (heterocedasticidade) Dae-Young (2023).

Adicionalmente, foi utilizando a técnica de combinação de Box-Cox e GARCH nos dados de temperatura das Bombas P1 e P2, com o objetivo de aprimorar a precisão dos modelos ao lidar com dados de séries temporais voláteis e com distribuições não normais, pois também segundo Dae-Young (2023), os métodos de GARCH e transformação Box-Cox auxiliam na redução de problemas relacionados a heterocedasticidade e não-normalidade nos resíduos, isso se os resíduos apresentarem correlação serial, heterocedasticidade e não-normalidade.

**Trabalhos relacionados/Estudo de caso**

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

O procedimento operacional requerido anteriormente, demandava uma série de atividades obrigatórias para melhor compreensão de algum desvio nos parâmetros, após seguidas todas as etapas de avaliação de desempenho conforme descrito pelo fabricante, era possível obtenção dos dados de temperatura:

* 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, como pode ser observado nas figuras 8 e 9 a seguir.

|  |  |
| --- | --- |
| Uma imagem contendo no interior, mesa, diferente, colorido  Descrição gerada automaticamente | Imagem de vídeo game  Descrição gerada automaticamente com confiança média |
| Figura 8. Termografia nos Cilindros de Elevação da Escavadeira  Fonte: Próprio autor | Figura 9. Termográfica com 5°C entre os pontos dos cilindros de elevação da lança  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).

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 e Proposta de solução**

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 didática.

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 identificados durante a análise das causas para melhorar o processo de Manutenção Preditiva, e eliminar a progressão dos danos internos no cilindro são apresentados a seguir na figura 13. Com base nesta avaliação, e com base nessa análise foi sugerido ações para melhorar o processo.

|  |
| --- |
| 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 modelo 395, primeiramente essa decisão foi com base no fato de que este modelo de ativo foi o mesmo 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.

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 | Texto  Descrição gerada automaticamente |
| Figura 15. Componentes monitorados  Fonte: Próprio autor | Figura 16. Estrutura de componentes  Fonte: Próprio autor |

*Nota: Em verde, o componente definido para análise neste estudo*

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

**Componentes/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 Bomba Principal 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 P1 e P2.**

Com base no descrito anteriormente, foram instalados os sensores nos componentes, conforme apresentado na figura 18, bem como a integração com a base de dados da Dynamox, posteriormente foi feita a integração dos dados disponibilizadas na Cloud para a análise dos dados de temperatura apresenta o comportamento apresentado na figura 18 a seguir.

|  |
| --- |
|  |
| 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 do Data Frame.

|  |
| --- |
|  |
| 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.

|  |
| --- |
|  |
| 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 11/09/2024 até 28/09/2024, compondo 8226 observações.

|  |
| --- |
|  |
| Figuras 21. Resumo dados temperatura Bomba P1 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 P1 e Bomba P2,

|  |
| --- |
|  |
| 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 P1, e uma assíntota a direita para os dados da Bomba P2.

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**

O teste ADF (Augmented Dickey-Fuller) examinou a hipótese nula de que uma série temporal é não-estacionária, complementados pelo teste KPSS (Kwiatkowski-Phillips-Schmidt-Shin), que considera a hipótese nula de estacionaridade. Conforme mostrado nas figuras 23a e b, todas as séries temporais são estacionárias, sem raiz unitária, em um nível de significância convencional (p < .05 ).

|  |  |
| --- | --- |
|  | |
| Figuras 23a. Valor-p Bomba P1  Fonte: Próprio Autor | Figuras 23b. Valor-p Bomba P2  Fonte: Próprio Autor |

A seguir, temos a figura 24, e testes que foram feitos com o objetivo de avaliar os resultados de métricas de performance/aderência destes modelos a série temporal estudada, algumas técnicas de AST&P (Advanced Statistical Techniques & Procedures) foram utilizadas e nesta figura temos os resultados apresentados em conjunto com técnicas mais avançadas de Machine Learning.

A definição para o modelo mais aderente para trabalhar com séries temporais, depende de vários fatores, incluindo a estrutura dos dados, a quantidade de dados disponíveis, a presença de tendências ou sazonalidades, entre outros.

Neste estudo, consideraremos os resultados dos modelos, quanto a sua aderência aos dados de ambos os componentes, a Bomba P1 e Bomba P2, e as métricas MAE, MSE, RMSE, AIC, e R².

Verifica a seguir nas figuras 24 e 25, estão grifados em cor laranja, os modelos apresentam maior aderência.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Modelo** | **MAE** | **MSE** | **RMSE** | **AIC** | **R2** | **MAE\_rank\_PumpP1** | **MSE\_rank\_PumpP1** | **R2\_rank\_PumpP1** | **AIC\_rank\_PumpP1** | **RMSE\_rank\_PumpP1** | **Total\_rank\_PumpP1** | | **RNNLSTMMainPumpP1** | **0.027568** | **0.002271** | **0.047658** | **-7935672683,0** | **0.951653** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** | **5.0** | | **RNNGRUMainPumpP1** | **2032892,0** | **13158910,0** | **3627521,0** | **5888750884,0** | **0.951507** | **2.0** | **2.0** | **2.0** | **4.0** | **2.0** | **12.0** | | **SVR\_AjusMainPumpP1** | **3246271,0** | **26042580,0** | **5103193,0** | **-5575905823,0** | **0.896335** | **3.0** | **3.0** | **3.0** | **2.0** | **3.0** | **14.0** | | **SVRMainPumpP1** | 9105355,0 | 148330137,0 | 12179086,0 | 8231078975,0 | 0.450596 | 4.0 | 4.0 | 4.0 | 5.0 | 4.0 | 21.0 | | **RLMMainPumpP1** | 12301500,0 | 214539995,0 | 14647184,0 | -3349901458,0 | 0.205359 | 5.0 | 5.0 | 5.0 | 3.0 | 5.0 | 23.0 | | **SARIMAXMainPumpP1** | 13545804,0 | 284065765,0 | 16854251,0 | 45574007860,0 | -0.037039 | 6.0 | 6.0 | 6.0 | 7.0 | 6.0 | 31.0 | | **ARIMAMainPumpP1** | 22793603,0 | 729091912,0 | 27001702,0 | 21450197862,0 | -0.531660 | 7.0 | 7.0 | 7.0 | 6.0 | 7.0 | 34.0 | |
| Figuras 24. Resumo dados temperatura Bomba P1  Fonte: Próprio Autor |

Para a Bomba P1, o modelo LSTM se destaca como o melhor modelo devido ao desempenho superior em todas as métricas, especialmente MAE, MSE, RMSE e AIC. Mesmo caso para a Bomba P2 o LSTM obteve bons resultados, já o SVR, mesmo tendo sido utilizado técnicas de ajuste com melhores parâmetros (RandomizedSearchCV), não obteve resultados satisfatórios, se comparado ao GRU e LSTM.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Modelo** | **MAE** | **MSE** | **RMSE** | **AIC** | **R2** | **MAE\_rank\_PumpP2** | **MSE\_rank\_PumpP2** | **R2\_rank\_PumpP2** | **AIC\_rank\_PumpP2** | **RMSE\_rank\_PumpP2** | **Total\_rank\_PumpP2** | | **RNNLSTMMainPumpP2** | **0.013628** | **0.000441** | **0.020990** | **-10631836236,0** | **0.993651** | **1.0** | **1.0** | **1.0** | **1.0** | **1.0** | **5.0** | | **RNNGRUMainPumpP2** | **1034305,0** | **2007230,0** | **1416767,0** | **2797466256,0** | **0.993359** | **2.0** | **2.0** | **2.0** | **4.0** | **2.0** | **12.0** | | **SVR\_AjusMainPumpP2** | 2493940,0 | 13837515,0 | 3719881,0 | -6564268128,0 | 0.945406 | 3.0 | 3.0 | 3.0 | 2.0 | 3.0 | 14.0 | | **SVRMainPumpP2** | 8643995,0 | 133752630,0 | 11565147,0 | 8060802907,0 | 0.555612 | 4.0 | 4.0 | 4.0 | 5.0 | 4.0 | 21.0 | | **RLMMainPumpP2** | 12402723,0 | 194681145,0 | 13952819,0 | -3509782559,0 | 0.353179 | 6.0 | 6.0 | 5.0 | 3.0 | 6.0 | 26.0 | | **ARIMAMainPumpP2** | 12272625,0 | 185629798,0 | 13624603,0 | 21399465381,0 | -0.000070 | 5.0 | 5.0 | 6.0 | 6.0 | 5.0 | 27.0 | | **SARIMAXMainPumpP2** | 13928457,0 | 326570126,0 | 18071251,0 | 26740511459,0 | -0.087035 | 7.0 | 7.0 | 7.0 | 7.0 | 7.0 | 35.0 | |
| Figuras 24a. Resumo dados temperatura Bomba P2  Fonte: Próprio Autor |

Para a Bomba P1, LSTM é claramente o melhor modelo, com os melhores resultados em todas as métricas, já o GRU também apresentou resultados satisfatórios, com MAE, MSE e R², bem como também se destacaram os resultados de RMSE, AIC. Para os dados desse componente, o SVR obteve resultados inferiores ao LSTM e GRU.

Ou seja, os modelos de GRU e LSTM obtiveram resultados bem mais aderentes aos dados, com os menores MAE e MSE e os R² mais altos.

Alinhado a isso, também o modelo SVR\_Ajustado com Randomized Search CV, como técnica de otimização de hiper parâmetros, maximizou a performance e apresentou melhoria no desempenho a ponto de ajustar mais a estes dados do que o modelo básico sem ajustes.

O sumário do modelo GARCH como visto na figura 25, aplicado à série temporal da Bomba P1 e Bomba P2, ajuda a entender mais o comportamento da volatilidade da média da série ao longo do tempo.

|  |
| --- |
| Tela de computador com texto preto sobre fundo branco  Descrição gerada automaticamente |
| Figuras 25. Sumário GARCH Bomba P1 e Bomba P2  Fonte: Próprio Autor |

No caso da Bomba P1, a série de dados apresentou média estável de 37.55, mas também alta volatilidade, sendo que o modelo GARCH mostra que a volatilidade da série é amplamente impactada por choques recentes (α₁ = 0.9348), enquanto a persistência de longo prazo da volatilidade (β₁ = 0.0505) foi menor.

Já para a Bomba P2, o modelo GARCH indica uma alta dependência de eventos recentes na volatilidade e uma média relativamente estável, também uma baixa persistência de longo prazo da volatilidade, o que pode ser entendido como sendo pelo fato de que os choques passados possam estar sendo rapidamente absorvidos pela série. A significância dos parâmetros sugere que o modelo está bem ajustado para capturar a dinâmica da volatilidade dessa série temporal.

Bomba P1 parece ter um ajuste ligeiramente melhor, com base no AIC e log-likelihood comparativamente a Bomba P2, com parâmetro de persistência de volatilidade (β) insignificante para Bomba P2, sugerindo que este modelo GARCH captura apenas os efeitos de choques recentes, sem persistência de volatilidade.

Ambos os modelos LSTM e GRU apresentam altos valores de α, indicando que os choques recentes têm grande impacto na volatilidade futura.

A variância incondicional (ω) é maior para Bomba P2, sugerindo que maior volatilidade não foi explicada no modelo.

Após a aplicação da transformação por Box Cox, e aplicado a análise de resíduos para Bomba P1 e Bomba P2 pelo teste Shapiro-Wilk, demonstraram p-valores extremamente baixos tanto para LSTM, GRU e SVR, confirmando a rejeição a H0 com distribuição não aderente à normalidade.

A Akaike Information Criterion (AIC) para Bomba P1 foi de (46914.7) e Bomba P2 (54891.3), auxiliou na comparação entre a qualidade dos modelos, considerando que neste caso quanto menor, melhor.

O modelo de Bomba P1 tem um AIC menor, sugerindo que ele pode ser um ajuste ligeiramente melhor para os dados do que o modelo para Bomba P2.

Já o resultado para o Log-Likelihood da Bomba P1 (-23453.3) e Bomba P2 (-27441.7), ou seja, mais uma métrica usada na avaliação de modelos, onde valores mais altos, ou seja, menos negativos indicam melhor ajuste.

Neste caso, Bomba P1 tem um log-likelihood maior, sugerindo um melhor ajuste comparado a Bomba P2.

**Resultados e Discussão**

A avaliação detalhada dos resultados permite identificar potenciais áreas de melhoria com uso de dados mássicos, coleta instantânea e análise estatística a fim de garantir o funcionamento eficiente e seguro das bombas Bomba P1 e Bomba P2 par aos próximos períodos.

Adicionalmente, a projeção com o modelo LSTM/GRU e SVR é apresentada na figura 25 e 26 utilizando os dados de temperatura par a Bomba P1. Já as figuras 27 e 28 foram feitas utilizando os dados de temperatura par a Bomba P2.

Em todos os casos a aderência aos dados reais foi forte, tendo sido utilizado a Transformação e Box Cox e Garch com o objetivo de ajustar os dados, observamos uma forte aderência aos modelos com dados de teste e treino.

|  |
| --- |
| **Gráfico  Descrição gerada automaticamente com confiança baixa** |
| Figuras 25. Temperatura Bomba P1 - LSTM/GRU e SVR - BOX COX  Fonte: Próprio Autor |
| **Gráfico, Histograma  Descrição gerada automaticamente** |
| Figuras 26. Temperatura Bomba P2 Modelo LSTM/GRU e SVR - BOX COX  Fonte: Próprio Autor |
| **Uma imagem contendo Histograma  Descrição gerada automaticamente** |
| Figuras 27. Temperatura Bomba P1 Modelo LSTM/GRU e SVR - GARCH  Fonte: Próprio Autor |

|  |
| --- |
| **Interface gráfica do usuário  Descrição gerada automaticamente com confiança média** |
| Figuras 28. Temperatura Bomba P2 Modelo LSTM/GRU e SVR - GARCH  Fonte: Próprio Autor |

Em todos os casos, os modelos baseados em redes neurais (LSTM e GRU) superaram os modelos tradicionais (SARIMAX, ARIMA e SVM), com o LSTM mostrando desempenho superior, especialmente para a Bomba P2, apresentando as melhores métricas e maior aderência aos dados, conforme mostrado na figura 29\_a e b.

As funções ACF (Autocorrelação) e PACF (Autocorrelação Parcial) dos resíduos dos modelos LSTM e GRU indicam que praticamente todos os lags estão dentro dos limites de confiança, sugerindo ausência de autocorrelação significativa. Isso demonstra que os modelos capturam bem os padrões temporais, sem deixar dependências seriais nos resíduos. Já o modelo SVR apresenta um comportamento diferente, com ACF mostrando alta autocorrelação nos primeiros lags, que diminui lentamente, e PACF confirmando essa correlação nos primeiros lags, especialmente no lag 1.

Neste ponto, os modelos LSTM e GRU mostraram-se mais adequados em ajustar-se aos dados, uma vez que seus resíduos não apresentaram autocorrelações significativas, sugerindo que os padrões temporais foram bem capturados. Em contraste, o modelo SVR apresentou autocorrelações persistentes nos resíduos, o que aponta para a necessidade de ajustes ou a consideração de outro modelo mais apropriado para captar as dinâmicas temporais.

|  |  |
| --- | --- |
|  |  |
| Figuras 29\_a. ACF e PACF Bomba P1  Fonte: Próprio Autor | Figuras 29\_b. ACF e PACF Bomba P2  Fonte: Próprio Autor |

Avançando na análise dos resultados dos modelos LSTM, GRU e SVR\_Ajustado, buscamos entender melhor o comportamento dos resíduos, tanto para a Bomba P1 quanto para a Bomba P2, agora adicionando um quadro comparativo para os testes de Shapiro-Wilk, Kolmogorov-Smirnov, Durbin-Watson e Breusch-Pagan. Esses testes são fundamentais para verificar a normalidade dos resíduos, a presença de autocorrelação e heterocedasticidade, garantindo que os modelos estejam adequadamente ajustados e que as conclusões sobre qual o modelo mais adequado para analise sejam robustas.

|  |  |
| --- | --- |
| **Teste** | **Análise** |
| Shapiro-Wilk (Teste de Normalidade) | Para a **Bomba P1**, todos os modelos (GRU, LSTM, SVM, SVR\_Ajus) têm p-valores extremamente baixos (na ordem de 10−32 10−32), o que significa que os resíduos não seguem uma distribuição normal.  Para a **Bomba P2**, ocorre o mesmo, com p-valores próximos a 10−3210^{-32}10−32, indicando não-normalidade dos resíduos. |
| Kolmogorov-Smirnov (Ajuste da Distribuição) | Para a **Bomba P1**, todos os modelos mostram p-valores muito baixos, especialmente o SVM e SVR\_Ajus (com um estatístico de 0.537 e p-valor essencialmente 0).  Para a **Bomba P2**, o modelo GRU também mostra p-valores baixos, indicando má adequação dos resíduos. |
| Durbin-Watson (Autocorrelação) | Para a **Bomba P1**:   * Os modelos GRU (2.94) e LSTM (2.74) estão próximos de 2, indicando baixa autocorrelação. * Os modelos SVM e SVR\_Ajus apresentam valores muito baixos (~0.26), sugerindo forte autocorrelação positiva.   Para a **Bomba P2**, o modelo GRU mostra um valor baixo (0.62), o que indica autocorrelação significativa. |
| Breusch-Pagan (Heterocedasticidade) | Para a **Bomba P1**, o modelo GRU tem um estatístico elevado (143.19) e um p-valor baixo, indicando heterocedasticidade. Os modelos SVM e SVR\_Ajus mostram estatísticas extremamente altas e p-valores próximos de zero, indicando forte heterocedasticidade.  Para a **Bomba P2**, o modelo GRU também apresenta alta heterocedasticidade com um estatístico de 320.09. |

Com base na análise dos resíduos, os modelos LSTM e GRU foram os melhores para a Bomba P1, com resultados mais favoráveis nos testes de Durbin-Watson (menor autocorrelação) e Breusch-Pagan (menor heterocedasticidade). No entanto, ambos enfrentam desafios relacionados à normalidade dos resíduos. Para a Bomba P2, o GRU apresentou algum nível de heterocedasticidade e autocorrelação, mas ainda assim teve um desempenho superior ao SVR\_Ajustado em termos de ajuste da distribuição.

Em resumo, o modelo LSTM mostrou-se mais consistente, com resíduos próximos de zero e menor dispersão, enquanto o SVR, apesar de uma média de resíduos próxima de zero, exibiu maior variabilidade e resíduos extremos, sugerindo a presença de outliers. O GRU teve desempenho intermediário entre o LSTM e o SVR.

Assim, mesmo após a análise dos resíduos, o LSTM demonstrou ser o modelo mais confiável para ambas as bombas, apresentando menos problemas de autocorrelação e heterocedasticidade.

|  |
| --- |
|  |
| Figuras 31. Teste dos resíduos para os modelos GRU, LSTM e SVR (Bomba P1)  Fonte: Próprio Autor |

|  |
| --- |
|  |
| Figuras 32. Teste dos resíduos para os modelos GRU, LSTM e SVR (Bomba P2)  Fonte: Próprio Autor |

**Avaliação de Overfitting (MSE e MAE para treino e teste)**

Para a Bomba P1, o LSTM apresenta uma boa generalização, com os valores de MSE e MAE menores no conjunto de teste em comparação com o treino indicando que o modelo não está sofrendo de overfitting e tem boa capacidade de prever os dados de teste.

O GRU também apresenta uma boa generalização, com valores de erro no conjunto de teste ligeiramente melhores do que no treino, como o MSE e MAE no teste são menores, o GRU parece ser ligeiramente superior ao LSTM para a Bomba P1.

O SVR apresenta erros maiores tanto no treino quanto no teste em comparação com o LSTM e o GRU, além disso, o erro no conjunto de teste é maior do que no conjunto de treino, sugerindo um possível overfitting ou um modelo mal ajustado.

Para os dados da Bomba P2, novamente, o LSTM apresenta um bom desempenho com valores de erro menores no conjunto de teste, sugerindo uma boa generalização sem overfitting.

O GRU tem o melhor desempenho em ambos os conjuntos de treino e teste, com os menores valores de MSE e MAE. Isso indica que o modelo está bem ajustado e generaliza melhor do que o LSTM e SVR para os dados da Bomba P2.

O SVR continua apresentando os piores resultados, com valores mais altos de MSE e MAE tanto no treino quanto no teste. Além disso, o erro de teste é ligeiramente superior ao erro de treino, sugerindo um ajuste fraco e possível overfitting

Os resultados obtidos para a amostra de dados coletados demonstraram diferentes comportamentos entre os modelos estudados. Os modelos de aprendizado de máquina, especificamente GRU e 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 mais eficazes.

**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 GRU e LSTM demonstraram um desempenho superior, evidenciado pelos menores valores de MAE e MSE e altos valores de R², indicando uma maior 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 melhores, posicionando-se como uma alternativa dentre os modelos utilizados no estudo. 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.

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.

Dae-Young Kim 2023. COVID-19 and Gun Violence: Keeping Unknown Shocks and Volatility in Perspective. Criminal Justice Review, © 2022 Georgia State University, Article Reuse Guidelines 48: 145-167.

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.