**Projeto de Pesquisa e Planejamento de Atividades**

|  |  |  |
| --- | --- | --- |
| **Aluno**: Cid Clay Aparecido Quirino | | **Data início curso**: 23/04/2023 |
| **Orientador**: Eder Costa Cassettari | | **Defesa em:**  12/2024 |
| **Curso**: MBA Data Science e Analytics | **Modalidade**: Distância | Turma: 231 |

1. **Título do projeto**

Otimização da condição de equipamentos de mineração, usando manutenção preditiva e IoT com Aprendizado de máquina.

1. **Introdução**

Nos últimos anos, diversos são os artigos e estudos demonstrando os avanços no monitoramento de equipamentos industriais de grande porte, tendo como principais objetivos a redução a utilização de seres humanos devido aos riscos inerentes, dificuldade de acesso aos locais de execução e redução na parada dos equipamentos com a antecipação de falhas, Tian, Y. M. et al. (1992) aponta que por outro lado, a produtividade e utilização dos equipamentos mecânicos continua a aumentar, 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 de grande porte 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), descreve que 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. Para avaliar a condição do equipamento, a manutenção preditiva utiliza testes não destrutivos, sensoriamento e coleta de dados de parâmetros tais como temperatura e vibração, análise de nível sonoro e outros testes em tempo real.

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.

Quanto aos ativos utilizados em aplicações de mineração, e possível identificar diferentes níveis de monitoramento, visto que alguns sistemas e componentes possuem maior cobertura por sensoriamento, tais como os componentes principais: Motores de combustão interna, Transmissão, Pneus e Comandos Finais, já os componentes menores, como cilindros, bombas e motores hidráulicos, são poucos os pontos de monitoramento, e quando são identificados, a captura dos dados e o monitoramento é realizado de maneira indireta, como exemplo podemos citar o monitoramento em alguns modelos de maquinário, onde a temperatura dos componentes do sistema hidráulicos que é feita no tanque, utilizando este dado como sendo a temperatura do sistema e não do componente em si, o que pode gerar retardo na tomada de decisão em caso de variação.

Artur Skoczylas et al (2023) avalia que cada vez mais é necessário um sistema aprimorado que capture, gerencie e seja eficaz no apoio a decisão para o gestor da manutenção, segundo afirma Artur Skoczylas et al (2023), é fundamental que as empresas empenhem em ter ecossistemas com capacidade de centralizar os vários aspectos importantes dos ativos moveis, a fim de garantir o monitoramento continuo do estado das máquinas de maneira mais ampla, a fim de prever avarias e planejar trabalhos de reparação antecipadamente (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 comenta que a 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 devido aos elevados custos operacionais.

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, com a aplicação de sensores de temperatura e vibração, a fim de coletar parâmetros e identificar mudanças no comportamento desses parâmetros ao longo de uma série temporal, com isso definir por uma intervenção antes da ocorrência de falhas.

* 1. **Objetivo**

A principal oportunidade identificada neste estudo, foi a necessidade de avançar com os atuais processos de Gerenciamento de Monitoramento de Condições (CMMS) para pequenos componentes, utilizando para tal, a aplicação de sensores de temperatura e vibração, possibilitando com isso, a coleta em tempo real de parâmetros ao longo de uma série temporal, objetivando identificar variação nos padrões e definir por uma intervenção antes da ocorrência de falhas.

Com base nessa dificuldade, 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.

A justificativa deste estudo, está nas dificuldades identificadas nas máquinas moveis utilizadas em mineração, de se obter parâmetros preditivos para o monitoramento das condições em pequenos componentes, e com isso, melhorar a previsibilidade de intervenção antes das falhas.

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.

1. **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.

* 1. **Manutenção Preditiva, Preventiva e Corretiva**

Manutenção Preditiva:

Manutenção Preventiva:

Manutenção Corretiva:

* 1. **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, 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 1: Device design and network  Fonte: H. Aguirre-Jofre et al. |

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 2 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 2: 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 de objetos, pessoas, sistemas e recursos de informação interconectados, juntamente com serviços inteligentes para permitir que eles processem informações sobre o mundo físico e virtual e reajam”

Este trabalho teve como ponto de partida, um estudo prévio que obteve resultados promissores em termos de monitoramento preditivo, utilizando câmera termográfica, 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.

* 1. **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 3 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 3: 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 4, 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 4: Visão Arvore de componentes Nuvem de Dados  Fonte: Próprio Autor |

* 1. **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.

* 1. **Trabalhos relacionados**

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.

* 1. **Estudo de caso**

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 5 e 6 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 5. Temperatura do Fluido e Figura 6. 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 7, 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 7. 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 8).

|  |
| --- |
| Imagem de vídeo game  Descrição gerada automaticamente com confiança média |
| Figura 8. 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 9 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 9. Cilindro na Centro de Reforma  Fonte: Próprio autor |

As Figuras 10a, 10b e 10c 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 10a. Falha interna 1  Fonte: Próprio autor | Figura 10b. Falha interna 2  Fonte: Próprio autor | Figura 10c. Falha interna 3  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?

* 1. **Métodos**

As etapas deste trabalho são descritas a seguir no Figura 11, desde a análise inicial até a conclusão.

Figura 11 – EAP projeto produto/processo

Fonte: Próprio autor

* + 1. **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 12 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 12: Ishikawa: Análise possíveis causas  Fonte: Próprio autor |

* + 1. **Definição da tecnologia IoT**

Consequentemente, uma proposta de solução foi requerida, e esquematizada na Figura 13, 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 13: New Online Data Collect  Fonte: Próprio autor |

* + 1. **Definição do equipamento/componentes para teste/validação em campo**

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 14 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 14: Escavadeira 395 e componentes a ser monitorados  Fonte: Próprio autor |

Já para gerenciamento dos parâmetros, foi definido uma estrutura de Sistema/Componentes conforme apresentado na Figura 15.

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

|  |
| --- |
|  |
| Texto  Descrição gerada automaticamente |
| Figura 15: Estrutura de sistemas/Componentes 395  Fonte: Próprio autor  \*\* Em verde, o componente definido para análise neste estudo |

* + 1. **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 16 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 16. Componentes monitorados  Fonte: Próprio autor |

* + 1. **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 17, 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 17 a seguir:

|  |
| --- |
| Linha do tempo  Descrição gerada automaticamente com confiança média |
| Figura 17: 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 18 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 18: Estrutura dos dados de temperatura e vibração  Fonte: Próprio Autor |

Já na Figura 19 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 19: 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 20 e 21 a seguir para a temperatura da Bomba Principal 1 e 2.

Os dados analisados e apresentados no Figuras 20 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 21 e 22 a seguir, são de um período separadas pela bomba 1 e bomba 2,

|  |
| --- |
| Gráfico, Histograma  Descrição gerada automaticamente |
| Figuras 21 e 22: 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.

* + 1. **Modelagem matemática**

A seguir as Figuras 23 e 24, 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 23: Valor-p para dados Bomba 1  Fonte: Próprio Autor | Figuras 24: Valor-p para dados Bomba 2  Fonte: Próprio Autor |

A definição pelo 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 nos podemos, foi elaborado uma matriz com os resultados consolidados dos modelos, elaborado utilizando a linguagem python e o Google Colab. Nesta tabela foram coletados os valores de MAE, MSE e R2 com o objetivo de identificar quais os modelos apresentam maior aderência aos dados coletados.

|  |
| --- |
| Modelo MAE MSE R2  0 State Space Model 0.531199 0.441022 -4.962101  1 SARIMAX 0.167584 0.042398 -0.261092  2 RNN-GRU 0.048013 0.006052 0.819975  3 RNN-LSTM 0.047513 0.005925 0.823764  4 Regressão Linear Múltipla 7.096247 94.210878 0.179421  5 Media Movel 2.911595 17.707430 0.837380  6 Model AR 7.648000 112.699675 0.535001  7 ARIMA 25.587862 885.270578 -2.647532  8 SVR 5.601069 62.932659 0.451855  9 Exponential Smoothing 0.162775 0.041150 -0.223966  10 Correlação 2.413219 25.064584 0.769936 |
| Figuras 25: Resumo dados temperatura Bomba 1 e 2  Fonte: Próprio Autor |

1. **Resultados Esperados**

Com base nos resultados apresentados para a amostra de dados coletados, os modelos estudados anteriormente apresentaram diferentes comportamentos, sendo que os modelos de aprendizado de máquina RNN-GRU e RNN-LSTM apresentaram o melhor desempenho em termos de métricas de erro (MAE e MSE) e explicação da variância (R2), ou seja, estes modelos foram capazes de capturar padrões complexos nos dados temporais.

Muito próximo aos resultados dos modelos de aprendizado está o modelo de Média Móvel, que também obteve um bom desempenho, como esperado, com baixas métricas de erro e um R2 alto, sugerindo como um modelo simples como a média móvel pode ser eficaz na previsão de séries temporais em certos casos.

Por outro lado, o modelo ARIMA apresentou um desempenho relativamente ruim em comparação com demais, tendo resultado em altos valores de MAE e MSE e um R2 negativo, o que indica que o modelo ARIMA pode não ter sido capaz de capturar adequadamente a estrutura dos dados de temperatura para a Bomba 1.

Já o caso dos modelos de regressão linear múltipla e SVR, estes mostraram desempenho intermediário, com valores moderados de MAE, MSE e R2, podendo ser mais bem estudado para ajustar os dados aos requisites destes modelos.

Complementando a análise, os modelos de State Space Model, SARIMAX e Exponential Smoothing apresentaram um desempenho inferior em comparação com outros métodos, com valores relativamente altos de MAE e MSE e R2 negativos, indicando que estes modelos podem não ser adequados para capturar a complexidade dos dados temporais de temperatura para a Bomba 1.

Portanto, para o conjunto de dados específico em consideração, os modelos RNN-GRU e RNN-LSTM foram os mais adequados considerando o ajuste aos dados coletados, e podem ser os mais adequados para realizar previsões precisas. Em contrapartida, os tempos de processamento para os modelos foram avaliados, indicando a necessidade de aprofundamento para entender o impacto deste tempo, quando na análise dos dados contínuos, o que estará sendo feito após a instalação dos sensores, como descrito no cronograma a seguir.

1. **Cronograma de Atividades**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Atividades planejadas** | **Mês** | | | | | | | | | |
| **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| Instalação dos sensores no ativo em campo | **X** |  |  |  |  |  |  |  |  |  |
| Coleta dos dados e início da análise (tempo real) | **X** |  |  |  |  |  |  |  |  |  |
| Validação dos métodos de análise e projeção presentados na metodologia |  | **X** |  |  |  |  |  |  |  |  |
| Identificação de pontos de melhoria e ajuste nas funções |  |  | **X** |  |  |  |  |  |  |  |
| Projeção dos dados de temperatura e identificação de pontos de risco |  |  |  | **X** |  |  |  |  |  |  |
| Entregar resultados preliminares (Introdução + Material e Método + resultados preliminares) ao Pecege |  |  |  | **X** |  |  |  |  |  |  |
| Fechar resultados e iniciar o resumo |  |  |  |  | **X** |  |  |  |  |  |
| Fechar e enviar para análise do orientador |  |  |  |  | **X** |  |  |  |  |  |
| Entrega do TCC e Agendamento da Defesa |  |  |  |  |  | **X** |  |  |  |  |
| Entrega da versão final revisada |  |  |  |  |  | **X** |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |

Projeto de Pesquisa; Resultados Preliminares; Entrega do Trabalho de Conclusão de Curso; Entrega da Apresentação da Defesa

1. **Referências Bibliográficas**

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.

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/>

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.