



UFOP

Universidade Federal
de Ouro Preto



Instituto de Ciências
Exatas e Aplicadas

Curso de Especialização em Ciência de Dados

Instituto de Ciências Exatas e Aplicadas

Universidade Federal de Ouro Preto

Bruno Rangel Araújo Vasconcelos

Gustavo Almeida De Souza Goncalves

Hylston Silva Lopes

**Estudo de Falhas no Sistema de Filtragem de Óleo Refrigerante de um
Laminador de Tiras a Frio Utilizando a Regressão Logística Como Ferramenta
de Auxílio à Decisão**

João Monlevade - MG

Fevereiro de 2024

Bruno Rangel Araújo Vasconcelos
Gustavo Almeida De Souza Goncalves
Hylston Silva Lopes

Trabalho de conclusão de curso apresentado ao curso de especialização em Ciência de Dados do Instituto de Ciências Exatas e Aplicadas da Universidade Federal de Ouro Preto, como parte dos requisitos necessários para a obtenção do título de Especialista em Ciência de Dados.

Orientador: Prof. Dr. Harlei Miguel de Arruda Leite

João Monlevade - MG

Fevereiro de 2024

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

L864e Lopes, Hylston Silva.

Estudo de falhas no sistema de filtragem de óleo refrigerante de um laminador de tiras a frio utilizando a regressão logística como ferramenta de auxílio à decisão. [manuscrito] / Hylston Silva Lopes. Bruno Rangel Araújo Vasconcelos. Gustavo Almeida de Souza Gonçalves. - 2024.
39 f.: il.: color., tab..

Orientador: Prof. Dr. Harlei Miguel de Arruda Leite.

Produção Científica (Especialização). Universidade Federal de Ouro Preto. Departamento de Engenharia de Produção.

1. Análise de regressão. 2. Confiabilidade (Engenharia). 3. Indústria 4.0. 4. Laminação (Metalurgia). 5. Processamento eletrônico de dados - Probabilidades. 6. Siderurgia. I. Gonçalves, Gustavo Almeida de Souza. II. Vasconcelos, Bruno Rangel Araújo. III. Leite, Harlei Miguel de Arruda. IV. Universidade Federal de Ouro Preto. V. Título.

CDU 519.2:004

Bibliotecário(a) Responsável: Flavia Reis - CRB6-2431



FOLHA DE APROVAÇÃO

Bruno Rangel Araújo Vasconcelos
Gustavo Almeida de Souza Gonçalves
Hylston Silva Lopes

Estudo de Falhas no Sistema de Filtragem de Óleo Refrigerante de um Laminador de Tiras a Frio Utilizando a Regressão Logística Como Ferramenta de Auxílio à Decisão

Trabalho de conclusão de curso apresentado ao curso de Especialização em Ciência de Dados da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de Especialista em Ciência de Dados

Aprovada em 27 de fevereiro de 2024

Membros da banca

Dr. Harlei Miguel de Arruda Leite - Orientador - Instituto Tecnológico de Aeronáutica
Dr. Thiago Augusto de Oliveira Silva - Universidade Federal de Ouro Preto
Me. Gustavo Costa Ávila - Aperam South America

Harlei Miguel de Arruda Leite, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 05/03/2024



Documento assinado eletronicamente por **Thiago Augusto de Oliveira Silva, PROFESSOR DE MAGISTERIO SUPERIOR**, em 11/03/2024, às 11:31, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0676988** e o código CRC **83290728**.

Agradecimentos

A Deus, por nossas vidas, e por nos permitir superar os desafios encontrados ao longo da realização deste curso.

A nossas famílias, por todo apoio durante o período do curso.

À Aperam South America, por proporcionar esse desenvolvimento em nossas carreiras e vidas.

À UFOP e seus professores, pela qualidade dos ensinamentos e atenção em todos os aspectos deste curso.

Ao professor Prof. Dr. Harlei Miguel de Arruda Leite por ter sido nossa orientador e por toda dedicação e paciência.

Aos colegas da turma de especialização em Ciência de Dados, pela sinergia e comprometimento que permitiu ao nosso grupo obter ótimos resultados.

Agradecemos a todos que, direta ou indiretamente, contribuíram para que este trabalho fosse concluído.

Resumo

O processo de laminação a frio em siderúrgicas envolve o uso de óleo como agente de resfriamento, que pode ser contaminado com o tempo. A filtragem regular do óleo é essencial para manter a eficiência e a longevidade do processo de laminação. Este estudo tem como objetivo desenvolver um modelo de regressão logística para prever a probabilidade de falha do sistema de filtragem de óleo de uma usina de laminação a frio de aço inoxidável. O estudo começa com uma descrição detalhada do sistema de filtragem de óleo, seus componentes e os fatores que podem contribuir para sua falha. Em seguida, o estudo coleta dados históricos sobre o sistema de filtragem, incluindo variáveis como queda de pressão, taxa de fluxo, temperatura e registros de manutenção. O estudo aplica a análise de regressão logística aos dados coletados para identificar as variáveis significativas que podem prever a falha do filtro. O modelo de regressão logística é então treinado e validado, e seu desempenho é avaliado usando métricas como exatidão, precisão, recuperação e pontuação. Análogo a isso, são usadas análises exploratórias de dados, por meio de software estatístico para certificar a ferramenta aplicada. O estudo conclui que o modelo de regressão logística pode prever a falha do filtro com 72,73% de precisão e acurácia de 87,16% e pode ser usado como uma ferramenta de apoio à decisão para a operação e manutenção do equipamento.

Palavras-chaves: Sistema de Óleo; Confiabilidade de Ativos; Indústria 4.0; Data Science Aplicado à Indústria; Análise de Depreciação.

Abstract

The cold rolling process in steel mills involves the use of oil as a cooling agent, which can become contaminated over time. Regular oil filtration is essential to maintain the efficiency and longevity of the rolling process. This study aims to develop a logistic regression model to predict the probability of failure of the oil filtration system of a stainless steel cold rolling mill. The study begins with a detailed description of the oil filtration system, its components and the factors that can contribute to its failure. Next, the study collects historical data on the filtration system, including variables such as pressure drop, flow rate, temperature and maintenance records. The study applies logistic regression analysis to the data collected to identify the significant variables that can predict filter failure. The logistic regression model is then trained and validated, and its performance is evaluated using metrics such as accuracy, precision, recovery and score. Analogous to this, exploratory data analysis using statistical software is used to certify the tool applied. The study concludes that the logistic regression model can predict filter failure with 72,73% of precision and 87,16% of accuracy and can be used as a decision support tool for equipment operation and maintenance.

Keywords: Oil System; Asset Reliability; Industry 4.0; Data Science Applied to Industry; Depreciation Analysis.

Lista de ilustrações

Figura 1 - Gráfico de Pareto das Falhas.....	12
Figura 2 – Laminador de bobinas Sendzimir Nº 3.	13
Figura 3 – Gráficos de Perda por Tempo e Frequência.....	14
Figura 4 – Esquema de Filtragem de Óleo.	14
Figura 5 – Esquema de um Laminador Sendzimir.....	16
Figura 6 – Curva Teórica da Regressão Logística.	19
Figura 7 – Conversão dos dados.	26
Figura 8 – Index e Drop da variável de fluxo de entrada.	26
Figura 9 – Características e alvo.....	27
Figura 10 – Amostras de Treinamento e Teste.	28
Figura 11 – Regressão logística - normalização, modelo e treino.	28
Figura 12 – Previsões e acurácia do modelo.....	29
Figura 13 – Matriz de Confusão.	29
Figura 14 – Precisão, Recall e F1 – Score.	30
Figura 15 – Fluxo de óleo – Resumo.	30
Figura 16 – Carta de Controle.....	31
Figura 17 – Decomposição das Séries Temporais.	31
Figura 18 – Análise por Componente.....	32
Figura 19 – Análise Sazonal.	32
Figura 20 – Análise de Capabilidade.....	33

Lista de tabelas

Tabela 1 – Variáveis do processo.....	25
---------------------------------------	----

Lista de abreviaturas e siglas

ABNT – Associação Brasileira de Normas Técnicas

COLAB – *Colaboratory* - Plataforma de *Data Science* do *Google*

LB3 – Laminador de Bobinas Número 3

MINITAB – *Software* de Análises Estatísticas

PIMS - *Plant Information Management System*

POWER BI – *Software* de Análises de Dados da *Microsoft*

Sumário

1 INTRODUÇÃO	11
1.1 Objetivo geral	12
1.2 Objetivos específicos	12
1.3 Justificativa	12
1.4 Contextualização do problema de pesquisa	13
2 REVISÃO BIBLIOGRÁFICA	15
2.1 Laminação a frio.....	15
2.2 Laminador Sendzimir.....	16
2.3 Filtragem de óleo de laminação	17
2.4 Análise de regressão logística	19
2.5 Estatística descritiva	21
3 METODOLOGIA	24
3.1 Construção da base de dados	24
3.2 Pré processamento de dados	25
3.3 Criação de tabela resposta, característica e alvo	26
3.4 Partição da base de dados	27
4 RESULTADOS	28
4.1 Avaliação da regressão logística.....	28
4.2 Análise descritiva dos dados	30
5 CONSIDERAÇÕES FINAIS	35
Referências.....	37

1 INTRODUÇÃO

A indústria siderúrgica enfrenta constantemente o desafio de manter a eficiência e a qualidade dos processos de laminação, que são essenciais para a produção de materiais metálicos de alto desempenho. Entre os diversos aspectos críticos que influenciam diretamente na qualidade do produto final e na produtividade da linha de laminação, a qualidade do óleo de laminação desempenha um papel fundamental.

O óleo de laminação é empregado em larga escala como lubrificante e refrigerante durante o processo de laminação, reduzindo o atrito entre o material e os rolos, dissipando o calor gerado e protegendo as superfícies metálicas contra a corrosão. No entanto, ao longo do tempo de operação, o óleo é suscetível à contaminação por partículas metálicas, resíduos de processo e outros contaminantes, comprometendo sua eficácia e, conseqüentemente, afetando a qualidade do produto final. (Brown, C., & White, D. 2021).

Neste contexto, a filtragem do óleo de laminação emerge como uma estratégia crucial para manter a integridade do fluido lubrificante/refrigerante, prolongando sua vida útil e garantindo a qualidade e a precisão dos processos de laminação. Este trabalho propõe explorar e analisar criticamente o sistema de filtragem para o óleo de laminação do Laminador de Bobinas Sendzimir (Tokyo Tech, 2005), investigando seus impactos na eficiência operacional, na qualidade do produto final e nos custos de manutenção.

Por meio desta pesquisa, pretende-se fornecer uma visão abrangente sobre a filtragem do óleo de processo, considerando diferentes variáveis operacionais e requisitos específicos da indústria siderúrgica de aços especiais. Ao final do trabalho, espera-se desenvolver uma análise por meio de *Data Science*, elaborando uma ferramenta de auxílio a decisão para o processo supramencionado.

Desta forma, este trabalho contribui para o avanço do conhecimento no campo da manutenção industrial e da otimização de processos, oferecendo insights valiosos para profissionais e pesquisadores envolvidos na gestão e operação de linhas de laminação a frio em indústrias siderúrgicas.

1.1 Objetivo geral

Desenvolver uma ferramenta de auxílio à decisão operacional, para reduzir as paradas devido a falha no sistema de filtragem de óleo do laminador de bobinas 3 da Aperam.

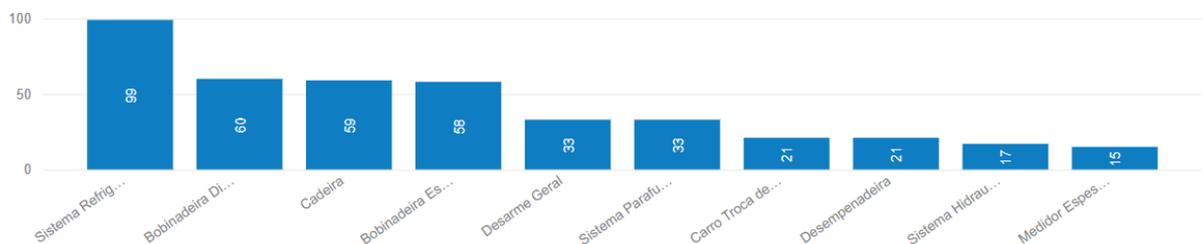
1.2 Objetivos específicos

- Entender a dinâmica das unidades de filtragem;
- Identificar as principais variáveis para montar uma estratégia de manutenção;
- Utilizar um modelo de *Data Science* de processo para sugerir os melhores métodos de operação do sistema de filtragem.

1.3 Justificativa

O sistema de filtragem de óleo de laminação do LB3 foi responsável por mais de 90 paradas no equipamento no ano de 2022, como mostra a Figura 1. Objetiva-se a utilização de *Data Science* para entender esses eventos e criar um modelo para auxílio a decisão, visando a redução no número de paradas deste sistema.

Figura 1 - Gráfico de Pareto das Falhas.



Fonte: Power BI – Aperam.

1.4 Contextualização do problema de pesquisa

O laminador a frio de inox 3 da Aperam, como mostra a Figura 2, utiliza óleo para refrigerar o processo de laminação das bobinas de aço. Este óleo é injetado na cabine de laminação por meio de bombas hidráulicas (com uma vazão controlada e predefinida por um modelo matemático) e posteriormente este óleo precisa ser filtrado antes de retornar à cabine. Caso haja falha no sistema de óleo refrigerante, pode ocasionar defeitos no material e/ou ruptura de tira.

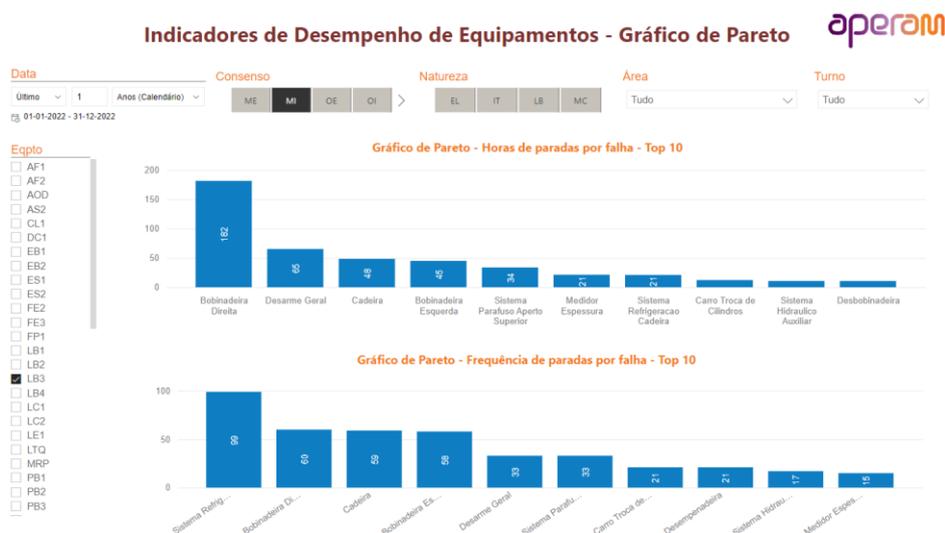
Figura 2 – Laminador de bobinas Sendzimir Nº 3.



Fonte: Aperam.

Contudo, algum distúrbio ou desbalanceamento na carga de entrada e saída das unidades de filtragem, tornam o sistema deficiente e conseqüentemente ocorre a parada no processo. Esses problemas acarretam na parada do equipamento, ocasionando perda no processo produtivo. Ao avaliar os gráficos de Pareto do LB3, identifica-se 02 famílias de perdas: Tempo e Frequência, como mostra a Figura 3.

Figura 3 – Gráficos de Perda por Tempo e Frequência.

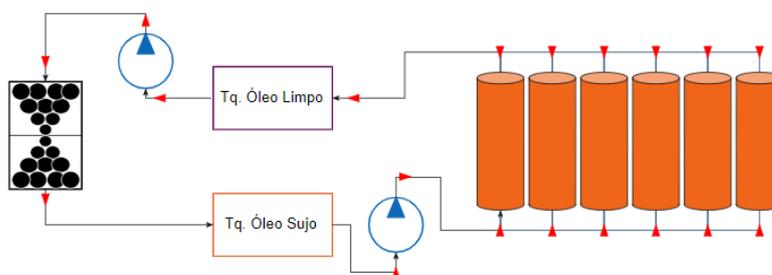


Fonte: Power BI – Aperam.

Quando se avalia pelo tempo de parada, temos “Bobinadeira Direita” como *Top* 1, e quando se avalia frequência de parada temos o “Sistema de Refrigeração” como Top 1. Em relação a Bobinadeira Direita, dia 28/07/2022 houve a quebra do eixo piramidal, gerando mais de 61 horas de parada (*Outlier*).

Em relação ao sistema de refrigeração não existem *outliers*, e as paradas são relacionadas aos tanques de óleo limpo e óleo sujo. Como uma parada de linha durante o processo pode ocasionar defeitos no material e até mesmo ruptura de tira, é decidido investigar o Pareto pela “Frequência”. Desse modo, o objeto de pesquisa será o “Estudo de Falhas no Sistema de Filtragem de Óleo Refrigerante de um Laminador de Tiras a Frio”. O processo é apresentado de forma esquemática pela Figura 4. Os triângulos vermelhos indicam os sensores do processo.

Figura 4 – Esquema de Filtragem de Óleo.



Fonte: Os autores.

2 REVISÃO BIBLIOGRÁFICA

2.1 Laminação a frio

A técnica de laminação a frio emergiu nos últimos anos da década de 1960 e ao longo da década de 1970 como uma resposta à crescente demanda por estruturas automotivas, de construção e outras estruturas de metal. Esse novo método foi concebido para minimizar a distorção, elevar a qualidade do material e otimizar a eficiência global em comparação com as técnicas tradicionais de laminação a quente (Oluwole, 2006).

Em síntese, a laminação é o processo de deformação plástica dos metais no qual o material é conduzido entre rolos. É o método mais comum de transformação mecânica de metais na prática devido à sua alta produtividade e controle dimensional preciso do produto final. Durante a laminação, o material é submetido a altas tensões compressivas, resultado da pressão exercida pelos rolos, e a tensões cisalhantes superficiais, geradas pela fricção entre os rolos e o material. As forças de fricção também são responsáveis por tracionar o metal.

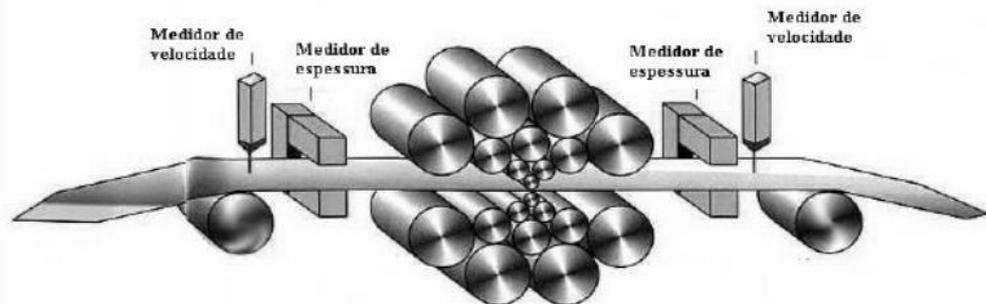
Os principais impulsionadores do avanço e aprimoramento da laminação a frio incluem progressos na tecnologia, ciência dos materiais e técnicas de fabricação mais sofisticadas (Nanjiang, 2014). Por exemplo, a utilização de múltiplos rolos desempenhou um papel fundamental na obtenção de maiores velocidades de produção e na redução do consumo de energia associado ao processo, além de viabilizar o processamento de materiais especiais.

Entretanto, apesar de suas numerosas vantagens, a laminação a frio enfrenta desafios como custos de produção elevados, limitações de equipamentos e o potencial surgimento de defeitos (Gautham, 2004). Para superar tais desafios, uma variedade de esforços contínuos de pesquisa e desenvolvimento está sendo empreendida para aprimorar o desempenho e a eficiência dos processos de laminação a frio.

2.2 Laminador Sendzimir

O laminador *Sendzimir* é uma máquina altamente versátil empregada na produção contínua de produtos metálicos. Esse tipo específico de laminador foi concebido por Henry Sendzimir, um químico e metalúrgico polonês, em 1857 (Bulgin, 1988), e desde então tem sido aprimorado e adaptado para atender às demandas das condições industriais modernas. Representado na Figura 5, o laminador *Sendzimir* é comumente utilizado na fabricação de chapas finas ou de grande diâmetro de diversos metais, incluindo aços inoxidáveis e aços elétricos.

Figura 5 – Esquema de um Laminador Sendzimir.



Fonte: Aperam.

A principal vantagem do laminador *Sendzimir* em comparação com sistemas anteriores, como os laminadores *Oliver* e *MacLaren*, reside na sua capacidade de alcançar alta precisão e elevada velocidade de produção (Tokyo Tech, 2005). A adoção de um motor elétrico, em contraste com a energia manual ou a vapor, aumentou ainda mais a eficiência e a adaptabilidade do laminador *Sendzimir*.

Recentemente, o laminador *Sendzimir* tem sido empregado na produção de materiais avançados, como aço de alta resistência, ligas de alumínio e superligas (Iyer & Smith, 2005). Além disso, tem sido utilizado na fabricação de materiais resistentes à corrosão para aplicações em engenharia marítima e offshore (Jadhav, 2013).

O futuro do laminador *Sendzimir* é promissor, à medida que os avanços na metalurgia, ciência dos materiais e engenharia assistida por computador continuam a expandir os limites dessa tecnologia. Isso inclui o desenvolvimento de laminadores *Sendzimir* automatizados, capazes de ajuste automático e detecção de falhas (Araújo et al., 2017).

2.3 Filtragem de óleo de laminação

A laminação a frio desempenha um papel crucial na indústria siderúrgica, onde a qualidade do óleo empregado desempenha uma função fundamental. A eficiente filtragem do óleo é crucial para assegurar a integridade do produto final e prolongar a vida útil dos equipamentos. Nesse contexto, vários estudos têm investigado diversas técnicas e estratégias de filtragem para otimizar o processo de laminação a frio.

Uma estratégia comum é a adoção de sistemas de filtragem de alta eficiência, como os filtros de cartucho. De acordo com *Smith et al. (2018)*, esses filtros têm a capacidade de eliminar partículas finas de óleo, contribuindo para reduzir o desgaste dos cilindros e melhorar a qualidade da superfície do material laminado.

Também pode-se avaliar a filtração centrífuga, que tem sido explorada como uma alternativa viável. Segundo *Jones (2020)*, esse método utiliza a força centrífuga para separar as impurezas do óleo, proporcionando uma limpeza mais completa e eficaz.

Outra estratégia promissora é a implementação de sistemas de monitoramento em tempo real. Conforme ressaltado por *Brown e Miller (2019)*, a utilização de sensores e dispositivos de monitoramento permite uma avaliação contínua da qualidade do óleo, possibilitando intervenções rápidas em caso de anomalias e garantindo um processo de laminação mais estável e confiável.

É relevante destacar também a importância dos estudos sobre reciclagem e regeneração do óleo de laminação. Segundo *Garcia et al. (2021)*, a recuperação do óleo usado por meio de processos de filtragem e purificação pode representar uma alternativa sustentável e econômica, reduzindo os custos operacionais e minimizando o impacto ambiental.

No caso específico do laminador *Sendzimir 3*, objeto de pesquisa deste estudo, são utilizados seis cartuchos de filtragem de óleo de processo, os quais, quando saturados, passam por uma retro lavagem. Esse procedimento, também conhecido como *retroflushing*, consiste em uma técnica empregada em sistemas de filtragem de óleo de laminação com o objetivo de remover as impurezas retidas nos elementos filtrantes. Essa etapa é crucial para manter a eficiência da filtragem,

prolongar a vida útil dos elementos filtrantes e garantir a integridade do óleo utilizado no processo de laminação a frio.

Um problema comum em sistemas de filtragem de óleo de processo de laminação é a contaminação do óleo por partículas sólidas, como poeira e resíduos metálicos, que podem resultar em danos aos componentes do equipamento de laminação e afetar a qualidade do produto final (Smith et al., 2018). Além disso, a formação de emulsões e a presença de água no óleo são questões críticas que podem levar à corrosão dos componentes e à deterioração das propriedades lubrificantes do óleo (Jones, 2020).

Outro desafio significativo enfrentado pelos sistemas de filtragem de óleo de laminação é a obstrução dos elementos filtrantes devido ao acúmulo de impurezas ao longo do tempo. Esse problema reduz a eficiência da filtragem e aumenta a resistência ao fluxo de óleo, resultando em uma menor capacidade de remoção de partículas e maior consumo de energia (Brown & Miller, 2019). Além disso, a obstrução dos filtros pode causar pressões elevadas no sistema, aumentando o risco de vazamentos e falhas nos componentes.

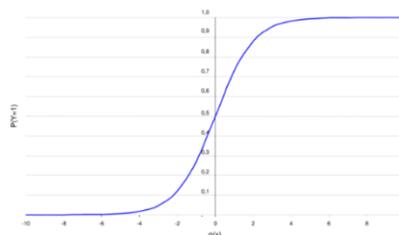
A degradação do óleo ao longo do tempo também é uma preocupação importante em sistemas de filtragem de óleo de laminação. A exposição a altas temperaturas e pressões, juntamente com a presença de contaminantes, pode levar à oxidação do óleo e à formação de produtos de decomposição prejudiciais, comprometendo suas propriedades lubrificantes e de resfriamento (Garcia et al., 2021).

Em resumo, os sistemas de filtragem de óleo de laminação enfrentam uma série de desafios, incluindo contaminação sólida, presença de água, obstrução dos filtros e degradação do óleo. Nesse contexto, a aplicação de ferramentas de apoio à decisão para as equipes de operação e manutenção deste processo torna-se cada vez mais necessária para a competitividade do negócio.

2.4 Análise de regressão logística

O método adotado neste estudo é a regressão logística, uma técnica estatística empregada para modelar a probabilidade de ocorrência de um evento, comumente utilizada em problemas de classificação binária. Nessa abordagem, a variável dependente é binária, representando dois possíveis resultados, enquanto as variáveis independentes podem ser contínuas ou categóricas. O modelo estima as chances de o evento ocorrer por meio de uma função logística, que transforma a soma ponderada das variáveis independentes em uma probabilidade entre 0 e 1. A interpretação dos coeficientes no modelo de regressão logística é realizada em termos de odds, fornecendo insights sobre a direção e magnitude do efeito de cada variável independente sobre a variável dependente (HOSMER Jr; LEMESHOW; STURDIVANT, 2013). Essa metodologia é amplamente aplicada em diversas áreas, como medicina, ciências sociais, engenharia e marketing. A representação teórica da curva de regressão logística pode ser observada na Figura 6:

Figura 6 – Curva Teórica da Regressão Logística.



Fonte: Os autores.

A Regressão Logística é um método estatístico amplamente empregado para analisar dados e prever resultados em problemas de classificação binária, onde a variável dependente é categórica e possui apenas dois possíveis valores (Hosmer et al., 2013). Essa técnica representa uma extensão da regressão linear, mas em vez de prever um valor contínuo, ela estima a probabilidade de ocorrência de um evento (Pedregosa et al., 2011).

A Regressão Logística foi inicialmente introduzida por Joseph Berkson em 1944, em um artigo intitulado "Are there 'Pure' Chance Experiments?" (Berkson, 1944). No entanto, o método ganhou destaque após o trabalho de David Cox em 1958, que desenvolveu a teoria matemática subjacente à Regressão Logística e a

aplicou em estudos epidemiológicos (Cox, 1958). Desde então, a Regressão Logística tem sido amplamente adotada em diversas áreas, como medicina, engenharia, ciências sociais e marketing (Hosmer et al., 2013).

A Regressão Logística fundamenta-se na função logística, uma função matemática que transforma qualquer valor real em um valor entre 0 e 1. Essa função é definida como:

$$p(x) = \frac{1}{1 + e^{(-b_0 - b_1 * x)}}$$

A probabilidade de ocorrência de um evento $p(x)$ na regressão logística é determinada pela variável independente x , enquanto os parâmetros estimados a partir dos dados são b_0 e b_1 (Pedregosa et al., 2011). A estimativa da regressão logística é realizada utilizando o método de máxima verossimilhança, que visa encontrar os valores dos parâmetros b_0 e b_1 que maximizam a probabilidade dos dados observados (Hosmer et al., 2013). Os coeficientes de regressão gerados pela análise de regressão logística podem ser interpretados como indicadores do efeito de cada variável independente na probabilidade de ocorrência do evento (Pedregosa et al., 2011).

Existem duas formas predominantes de regressão logística: a regressão logística simples, que envolve apenas uma variável independente, e a regressão logística múltipla, que incorpora várias variáveis independentes (Hosmer et al., 2013). A regressão logística múltipla permite a investigação de interações entre as variáveis independentes e a avaliação do impacto de cada variável enquanto controla os efeitos das demais variáveis (Pedregosa et al., 2011).

A regressão logística apresenta algumas vantagens em relação a outros métodos de classificação, como a interpretação direta dos coeficientes de regressão e a capacidade de lidar com conjuntos de dados desequilibrados, onde uma das categorias da variável dependente é muito mais frequente do que a outra (Pedregosa et al., 2011). No entanto, a regressão logística também possui limitações, como a suposição de linearidade entre as variáveis independentes e a variável *logit*, além da necessidade de um número considerável de observações para estimativas precisas (Hosmer et al., 2013).

A aplicação da regressão logística é vasta, sendo empregada na previsão de doenças cardiovasculares (Hosmer et al., 2013), no domínio do marketing para prever a probabilidade de um cliente adquirir um produto (Pedregosa et al., 2011), e na engenharia para identificar falhas em sistemas complexos (Sankararaman & Mahapatra, 2013). Além disso, a regressão logística é frequentemente combinada com outros métodos estatísticos, como árvores de decisão e redes neurais, para melhorar a precisão dos modelos de classificação (Kuhn & Johnson, 2013).

2.5 Estatística descritiva

A estatística descritiva é um ramo da estatística que se concentra na análise e interpretação de dados para resumir suas características principais de forma concisa e informativa. Ela envolve a organização, resumo e apresentação dos dados de uma maneira que facilite a compreensão das principais tendências e padrões presentes nos mesmos.

A estatística descritiva é uma ferramenta essencial em áreas como ciências sociais, negócios, economia, saúde e muitas outras, pois fornece insights valiosos sobre os dados e ajuda na tomada de decisões informadas (MORETTIN; BUSSAB, 2017).

Para o trabalho, foram realizadas análises estatísticas dos dados, sendo elas: Mapa de Calor, Carta de Controle, Séries Temporais, Tendência e Análise de Capacidade. Por meio de tais análises é possível entender o comportamento dos dados e auxiliar às decisões do processo de laminação, visando a filtragem de óleo e o volume de óleo disponível para uso.

Abordando a decomposição de série temporal é uma técnica que divide uma série temporal em seus componentes principais: tendência, sazonalidade e erro. Essa técnica é útil para entender melhor o comportamento de uma série temporal e para melhorar as previsões. (MORETTIN; BUSSAB, 2017).

A decomposição de série temporal pode ser aplicada a qualquer variável, mas é mais comumente usada para variáveis econômicas, como vendas, preços e

produção. Neste caso, foi utilizada para verificar a vazão de óleo de refrigeração e sua flutuação no tempo.

O método mais comum de decomposição de série temporal é o método aditivo. Esse método assume que a série temporal pode ser representada pela seguinte equação:

$$y(t) = T(t) + S(t) + E(t)$$

Onde:

$y(t)$ é o valor da série temporal no período t

$T(t)$ é a tendência da série temporal no período t

$S(t)$ é o componente sazonal da série temporal no período t

$E(t)$ é o componente de erro da série temporal no período t

A tendência da série temporal é o movimento geral da série temporal ao longo do tempo. Ela pode ser representada por uma linha reta ou uma curva. O componente sazonal da série temporal é o padrão de variação da série temporal ao longo do tempo. Ele pode ser causado por fatores como estações do ano, feriados ou eventos recorrentes. O componente de erro da série temporal é a variabilidade da série temporal que não pode ser explicada pela tendência ou pela sazonalidade. (MORETTIN; BUSSAB, 2017).

A decomposição de série temporal é uma ferramenta poderosa que pode ser usada para entender melhor o comportamento de uma série temporal e para melhorar as previsões. A decomposição de série temporal tem várias vantagens, incluindo:

- Ela pode ajudar a entender melhor o comportamento de uma série temporal. A identificação da tendência e da sazonalidade pode ajudar a explicar os padrões de variação da série temporal.
- Ela pode melhorar as previsões. A identificação da tendência e da sazonalidade pode ajudar a desenvolver modelos de previsão mais precisos.

- Ela pode ser usada para identificar anomalias. A identificação de valores que estão fora da tendência sazonal pode ajudar a identificar anomalias na série temporal.

A decomposição de série temporal também tem algumas desvantagens, incluindo:

- Ela pode ser subjetiva. A identificação da tendência e da sazonalidade pode ser um processo subjetivo, dependendo do método usado.

Ela pode ser complexa. A decomposição de série temporal pode ser um processo complexo, especialmente para séries temporais com padrões de sazonalidade complexos. (MORETTIN; BUSSAB, 2017).

3 METODOLOGIA

A metodologia utilizada para o presente estudo foi a análise dos dados por meio de *Python* (utilizando o ambiente de desenvolvimento *Google Colab*) e análise estatística (usando o *software Minitab*). Para ilustrar o passo a passo da metodologia aplicada, serão percorridos nas seções subsequentes a análise de regressão logística, a construção da base de dados e sua tratativa, além de demais passos para a construção do modelo.

3.1 Construção da base de dados

Para a elaboração da base de dados do problema de pesquisa, inicialmente foram extraídos os dados de 1 ano de operação do LB3, acessados através do sistema corporativo industrial PIMS. As variáveis obtidas através do sistema são gravadas em uma base a cada 1 segundo.

Em uma análise inicial do volume de dados e pela característica do processo, é decidido trazer o horizonte para um mês de observação (operação). Com isso, temos uma base com 65526 eventos para serem avaliados. Em função do processo, tal avaliação é considerada válida.

As informações extraídas do sistema computadorizado PIMS são salvas em uma planilha eletrônica no formato *.csv*. As informações supramencionadas são divididas em 21 variáveis, como mostra a Tabela 1.

Tabela 1 – Variáveis do processo.

Variável	Descrição
LFI_LB3_COILLD	ID da bobina
LFI_LB3_MILLVEL_VALIDA	Nível de óleo
LFI_LB3_VAZAO_OLEO_LIM_INF	Vazão de óleo geral
LFI_LB3_STATUS_PROCESSO	Status do processo
LFI_LB3_ROLLOIL_TEMP	Temperatura do óleo
LFI_LB3_ROLLOIL_FLOW_ENTRY	Fluxo de entrada de óleo
LFI_LB3_OLEOREFTEMP	Referência de temperatura do óleo
LFI_LB3_OLEOREFPRESS	Referência de pressão do óleo
LFI_LB3_OLEOFILTRO1VAZ	Vazão de óleo do filtro 1
LFI_LB3_OLEOFILTRO2VAZ	Vazão de óleo do filtro 2
LFI_LB3_OLEOFILTRO3VAZ	Vazão de óleo do filtro 3
LFI_LB3_OLEOFILTRO4VAZ	Vazão de óleo do filtro 4
LFI_LB3_OLEOFILTRO5VAZ	Vazão de óleo do filtro 5
LFI_LB3_OLEOFILTRO6VAZ	Vazão de óleo do filtro 6
LFI_LB3_OLEOACTFLUXO	Fluxo de óleo na cadeira
LFI_LB3_MILLVEL_VALIDA.1	Validação do nível de óleo
LFI_LB3_BOMB16K1LIG	Status da bomba 1
LFI_LB3_BOMB16K2LIG	Status da bomba 2
LFI_LB3_BOMB16K3LIG	Status da bomba 3
LFI_LB3_BOMB16K4LIG	Status da bomba 4
LFI_LB3_BOMB16K5LIG	Status da bomba 5

Fonte: Os autores.

3.2 Pré processamento de dados

Após a construção dos dados, são adicionadas duas colunas (novas variáveis), denominadas: PARADAS e PARADAS_P. Essas variáveis checam se a base de dados está coerente. Caso os filtros estejam desligados (bombas off), o equipamento está parado (PARADAS). Para validar a informação é criada a variável PARADAS_P.

Com a base de dados carregada no *Google Colab* é feita a avaliação de dados faltantes. Esse é um passo para garantir que no momento de treino e teste não existam valores que possam destoar do padrão, criando assim *outliers*. Ao aplicar o comando do *Pandas* "`df.isnull().sum()`", não foi encontrado dado faltante, possibilitando o avanço no trabalho.

Como a regressão logística (mencionada anteriormente) utiliza de uma análise binária (zero ou um), é realizada a substituição dos termos "On" para "1" e os

termos “Off” para “0”, como mostra a Figura 7. Essa alteração é feita no *Data Frame* através do comando “`df = df.replace`”.

Figura 7 – Conversão dos dados.

LFI_LB3_MILLVEL_VALIDA.1	LFI_LB3_BOMB16K1LIG	LFI_LB3_BOMB16K2LIG	LFI_LB3_BOMB16K3LIG	LFI_LB3_BOMB16K4LIG	LFI_LB3_BOMB16K5LIG
1178.0	1	1	1	0	0
2797.0	1	1	1	0	0
2456.0	1	1	1	0	0
2115.0	1	1	1	0	0
1774.0	1	1	1	0	0

Fonte: Os autores.

Do mesmo modo, são alterados os termos “Processando” para “1” e “Parado” para “0”.

3.3 Criação de tabela resposta, característica e alvo

Como o objeto de estudo do trabalho é a filtragem de óleo do LB3, é isolada a variável “LFI_LB3_ROLLOIL_FLOW_ENTRY” em um novo *Data Frame*. A este é dado o nome “new_data”.

Após essa segregação, é realizado um *index* e um *drop* com a tabela criada. Os dois modelamentos são executados com os comandos do *Pandas*: “new_data.index” e “new_data.drop”, como mostra a Figura 8.

Figura 8 – Index e Drop da variável de fluxo de entrada.

	LFI_LB3_ROLLOIL_FLOW_ENTRY	LFI_LB3_ROLLOIL_FLOW_ENTRY
0	1100	1100
1	1100	1100
2	1100	1100
3	1100	1100
4	1100	1100
...
65521	1100	1100
65522	1100	1100
65523	1100	1100
65524	1100	1100
65525	1100	1100

65526 rows x 1 columns 65526 rows x 0 columns

Fonte: Os autores.

Após essa etapa, são definidas no *Phyton* as características e alvo a serem avaliados no modelo, de acordo com a Figura 9.

Figura 9 – Características e alvo.

```
features = (['LFI_LB3_ROLLOIL_FLOW_ENTRY', 'LFI_LB3_OLEOFILTRO6VAZ'])  
target = 'LFI_LB3_STATUS_PROCESSO'
```

Fonte: Os autores.

3.4 Partição da base de dados

A base de dados foi particionada em dois subconjuntos distintos, o conjunto de treinamento e o conjunto de teste, seguindo uma proporção de 60% e 40%, respectivamente. Esta separação tem por objetivo prevenir o sobre ajuste dos dados e permitir uma avaliação mais precisa do desempenho e da capacidade de generalização do modelo preditivo desenvolvido.

4 RESULTADOS

4.1 Avaliação da regressão logística

Para a aplicação da regressão logística, a base de dados foi dividida em “treino” e “teste”, como apresentado na Figura 10. O tamanho da base de teste é de 40% da base total.

Figura 10 – Amostras de Treinamento e Teste.

```
print(f'Tamanho do conjunto de treinamento: {len(X_train)} amostras')
print(f'Tamanho do conjunto de teste: {len(X_test)} amostras')
```

```
Tamanho do conjunto de treinamento: 39315 amostras
Tamanho do conjunto de teste: 26211 amostras
```

Fonte: Os Autores.

Após essa etapa, foi elaborado o modelo de regressão logística e realizada as predições, como mostra a Figura 11.

Figura 11 – Regressão logística - normalização, modelo e treino.

```
#Normalizar os dados.
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)
```

```
#Um modelo de regressão logística é criado com regularização L1
model = LogisticRegression(penalty='l1', solver='liblinear', C=1.0)
```

```
# Treinar o modelo
model.fit(X_train_scaled, y_train)
```

```
LogisticRegression
LogisticRegression(penalty='l1', solver='liblinear')
```

Fonte: Os Autores.

O modelo obteve uma acurácia de 87,16%, como mostra a Figura 12, onde para um processo como este, pode-se ser considerado satisfatório.

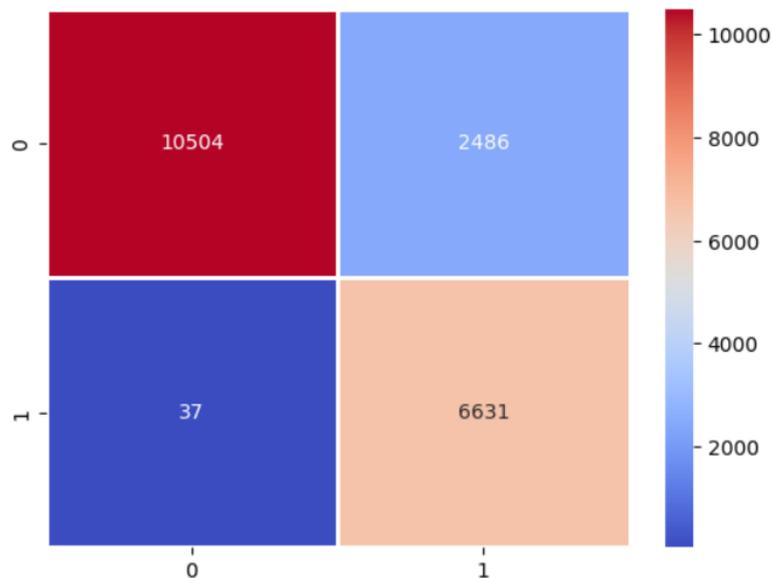
Figura 12 – Previsões e acurácia do modelo.

```
[ ] 1 # Avaliar o modelo
     2 accuracy = accuracy_score(y_test, predictions)
     3 print(f'Accuracy: {accuracy}')
     4
Accuracy: 0.8716553057279479
```

Fonte: Os Autores.

Para a certificação desse método, é analisada a Matriz de Confusão. Nela podemos observar o comportamento dos dados de acordo com a acuracidade do algoritmo desenvolvido, como mostra a Figura 13.

Figura 13 – Matriz de Confusão.



Fonte: Os Autores.

Essas informações obtidas através do algoritmo denotam uma taxa de acerto considerável, onde sua aplicação pode servir como uma ferramenta de auxílio a decisão. Pelas características do processo e das variáveis, a matriz de confusão confirma o objeto de estudo em questão. Após essa etapa, foi gerado o algoritmo para calcular a Precisão, Recall e F1 – Score, como mostra a Figura 14. Podemos avaliar uma precisão de 72,73%.

Figura 14 – Precisão, Recall e F1 – Score.

```
[ ] 1 from sklearn.metrics import precision_score, recall_score, f1_score
```

```
[ ] 1 # Calcular precisão, recall e F1-score
2 precision = precision_score(y_test, predictions)
3 recall = recall_score(y_test, predictions)
4 f1 = f1_score(y_test, predictions)
5
6 print(f'Precision: {precision}')
7 print(f'Recall: {recall}')
8 print(f'F1-score: {f1}')
```

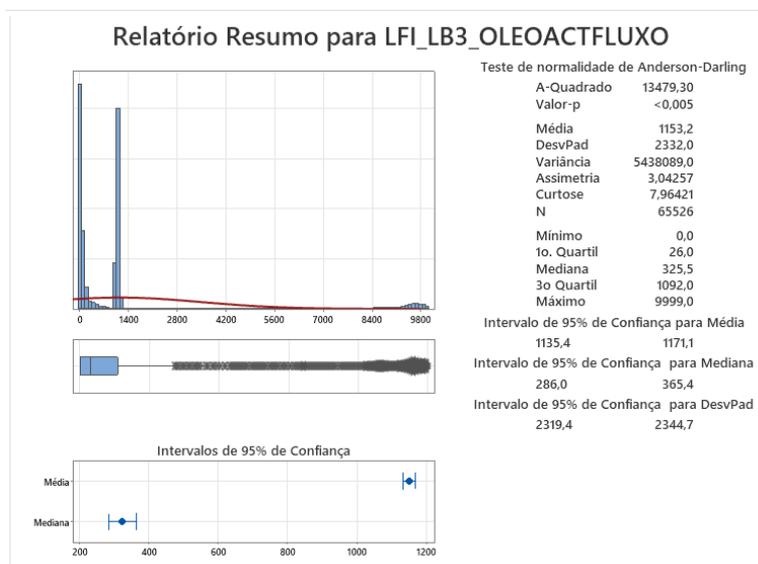
```
Precision: 0.7273225841833937
Recall: 0.9944511097780444
F1-score: 0.840164713335445
```

Fonte: Os Autores.

4.2 Análise descritiva dos dados

Em relação a variável de vazão de entrada, é possível entender o comportamento do sistema de filtragem. Essa variável denota o ritmo de processo e serve como *Input* para o processo de retro lavagem, em algumas etapas do processo de laminação. A Figura 15 apresenta o relatório resumo para os dados amostrados.

Figura 15 – Fluxo de óleo – Resumo.

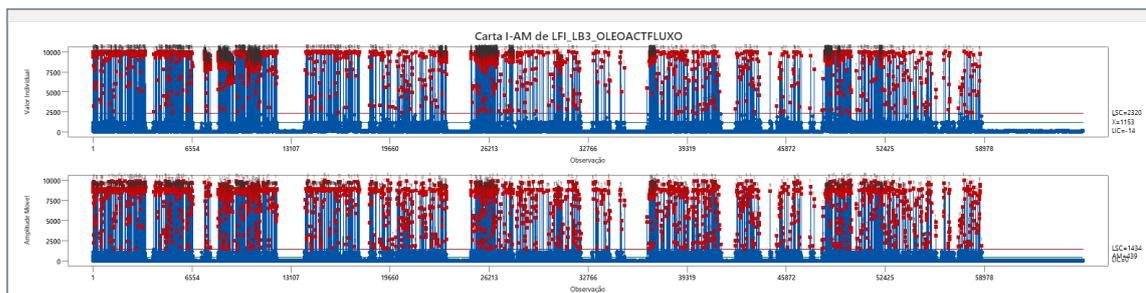


Fonte: Os Autores.

Nota-se que a maior parte da produção de aços exige uma vazão reduzida (até 1200 litros/min). Em uma avaliação por gráfico de contorno fica evidenciado que existe uma predominância nos dados de vazão nas faixas entre 3000 e 6000 litros/min. Contudo o problema maior reside nas maiores exigências de vazão do sistema de filtragem (9600 litros/min).

A análise da vazão pelo método de Carta de Controle, como mostra a Figura 16, apesar da grande variância, pode-se observar a estabilidade da variável dentro da faixa de 900 à 9600 litros/min.

Figura 16 – Carta de Controle.



Fonte: Os Autores.

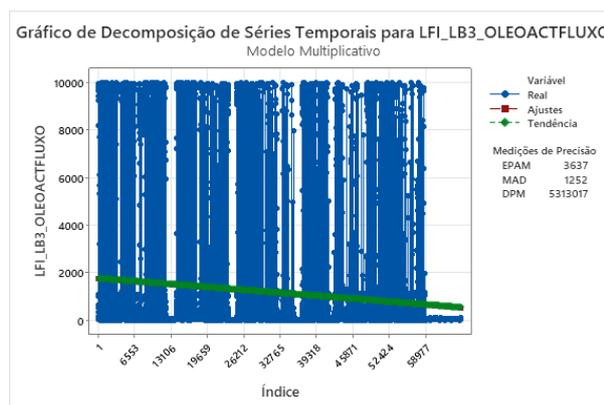
Durante as análises, obteve-se a seguinte equação de tendência ajustada.

Equação de Tendência Ajustada

$$Y_t = 1765,6 - 0,018691 \times t$$

Após essa etapa, pode-se avaliar a decomposição das séries temporais, como mostra a Figura 17.

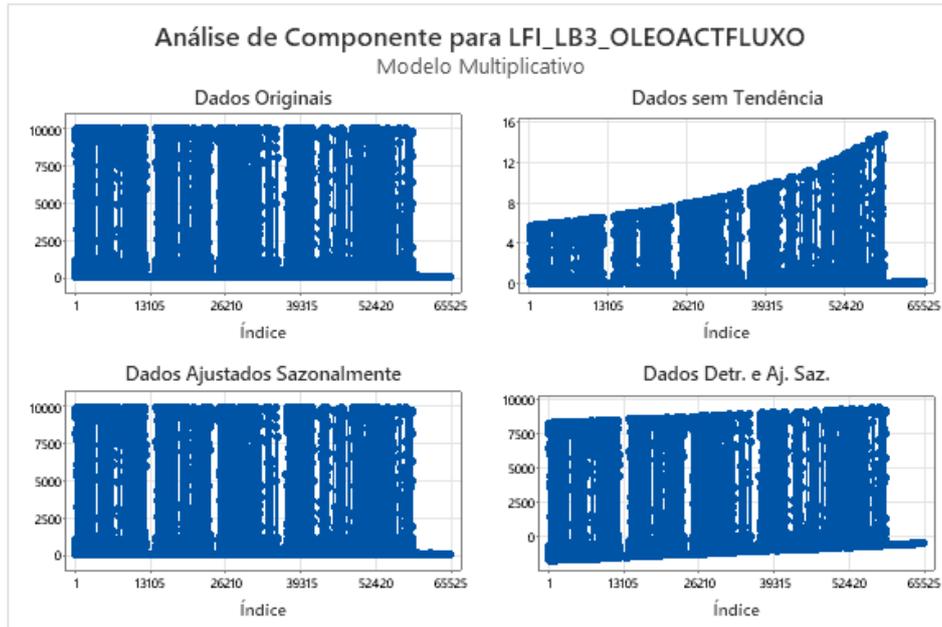
Figura 17 – Decomposição das Séries Temporais.



Fonte: Os Autores.

Fazendo a análise da variável de fluxo de óleo, podemos observar na Figura 18 o comportamento dos dados. Nota-se uma semelhança entre o comportamento das variáveis.

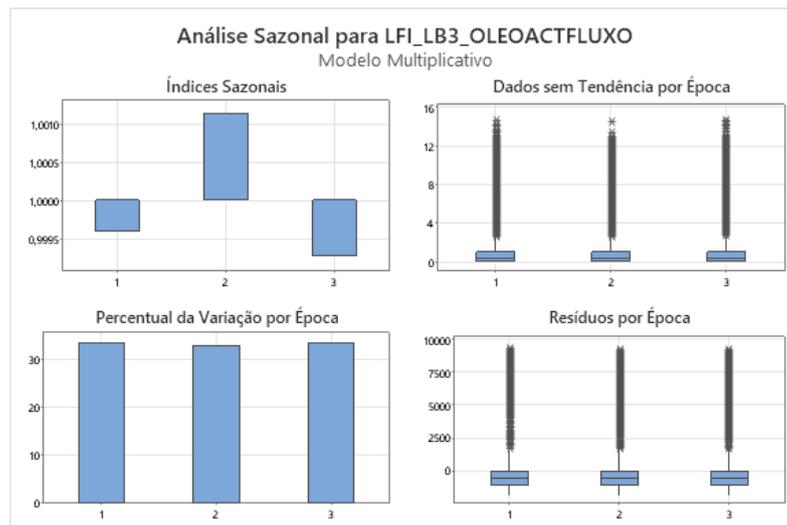
Figura 18 – Análise por Componente.



Fonte: Os Autores.

Já na Figura 19 é possível analisar a sazonalidade dos dados. Os dados são avaliados de forma sazonal, sem tendência e por época.

Figura 19 – Análise Sazonal.



Fonte: Os Autores.

Foi analisada a capacidade do processo de filtragem de óleo de refrigeração de fornecer no mínimo 8500 litros de óleo por minuto. Uma das principais características no teste de capacidade é o ZBench. O ZBench é uma medida da capacidade de um processo de produzir peças dentro de uma especificação. Ele é calculado como a distância entre a média do processo e o limite superior da especificação, dividido pelo desvio padrão do processo. A fórmula para calcular o zbench é a seguinte:

$$zbench = (\mu - USL) / \sigma$$

Onde:

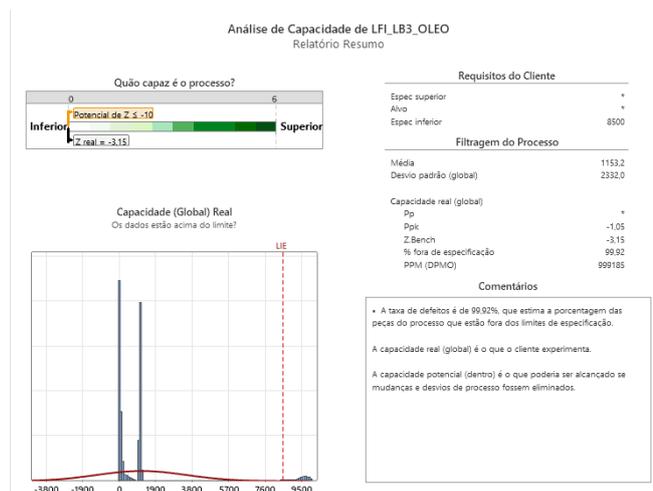
- μ é a média do processo;
- USL é o limite superior da especificação;
- σ é o desvio padrão do processo.

O zbench também pode ser usado para estimar a porcentagem da vazão de óleo que está dentro da especificação de 8500litros/min. A porcentagem dentro da especificação pode ser calculada usando a seguinte fórmula:

$$\% \text{ dentro da especificação} = 1 - (1.5 * zbench) ^ 2$$

Por exemplo, se o zbench for igual a 2, a porcentagem das vazões dentro da especificação será de 95,45%. No contexto de um teste de capacidade, como mostra a Figura 20, o zbench atinge - 3,15.

Figura 20 – Análise de Capacidade.



Fonte: Os Autores.

Como mencionado, o zbench obtido foi de -3,15. Um zbench de -3,15 significa que a média do processo está 3,15 desvios padrão abaixo do limite superior da especificação. Usando a fórmula para calcular a porcentagem de vazões de óleo dentro da especificação:

$$\% \text{ dentro da especificação} = 1 - (1.5 * z\text{bench}) ^ 2$$

$$\% \text{ dentro da especificação} = 1 - (1.5 * -3.15) ^ 2$$

$$\% \text{ dentro da especificação} = 1 - (-10.125)$$

$$\% \text{ dentro da especificação} = 0.89875$$

Portanto, apenas 89,875% das vazões do processo estarão dentro da especificação.

Um zbench de -3,15 é considerado um valor baixo, o que indica que o processo tem baixa capacidade de operar dentro da especificação com vazões mais elevadas. Esse valor sugere que o processo precisa de melhorias. Algumas ações que podem ser tomadas para melhorar a capacidade do processo incluem:

- Reduzir a variabilidade do processo;
- Aumentar a média do processo.

A redução da variabilidade do processo pode ser feita através de melhorias no processo, como treinamento de funcionários, uso de ferramentas e equipamentos adequados e controle de qualidade. O aumento da média do processo pode ser feito através de alterações nas receitas de laminação e mudanças na estratégia de manutenção dos ativos envolvidos (bombas, unidades de filtragem, válvulas, etc).

5 CONSIDERAÇÕES FINAIS

Ao longo deste trabalho, foi explorada a importância da filtragem eficaz do óleo em processos de laminação, destacando os benefícios significativos que essa prática proporciona à operação do laminador. Ficou claro que a qualidade do óleo utilizado desempenha um papel crucial na garantia da eficiência e da durabilidade dos equipamentos, bem como na qualidade dos produtos finais. Para isso, a filtragem deve estar sendo executada de forma eficaz.

Durante a análise, é perceptível que o uso da ciência de dados é de suma importância para a melhoria e garantia dos processos. Além disso, os principais parâmetros a serem considerados ao fazer uma análise acerca de um sistema de filtragem, como taxa de fluxo, entrada de óleo e bombeamento são imprescindíveis. Esses aspectos são essenciais para garantir um funcionamento eficiente do sistema de filtragem ao longo do tempo, minimizando os custos operacionais e maximizando o tempo de atividade do laminador.

É importante ressaltar que a implementação de práticas de manutenção preventiva e a realização de inspeções regulares são fundamentais para garantir o desempenho ideal do sistema de filtragem. A manutenção adequada não apenas prolonga a vida útil dos componentes do sistema, mas também ajuda a evitar paradas não programadas e reduzir o risco de danos aos equipamentos.

Pode-se observar que a aplicação da técnica de regressão logística obteve o sucesso esperado neste trabalho. O algoritmo pode prever a falha do filtro com 72,73% de precisão. A acurácia de 87,16% mostra que é possível a aplicação da ferramenta como auxílio a decisão, no processo de filtragem de óleo de laminação. Tal resultado dá-se avaliando as variáveis de fluxo de entrada de óleo, vazão do filtro 6 e o status do processo. Com essas variáveis de entrada e saída, é possível modelar o sistema de filtragem de óleo de processo e auxiliar na decisão da operação, sendo ela a mudança da receita de laminação, alteração do modelo matemático ou a parada do equipamento.

Por fim, enfatizamos a importância de investir em tecnologias de filtragem de alta qualidade e em processos de monitoramento contínuo para otimizar a eficiência operacional do laminador. Ao adotar uma abordagem proativa em relação à

manutenção e ao gerenciamento do sistema de filtragem, a operação do laminador pode alcançar níveis mais altos de produtividade, qualidade do produto e rentabilidade a longo prazo.

Como próximos passos, são sugeridas quatro ações. A redução da variabilidade do processo, a fim de garantir a estabilidade e controle operacional. Aumentar a média do processo, visando o atendimento da necessidade de filtragem para a laminação do material. Aprimorar a confiabilidade dos dados, para que as análises sejam fidedignas à realidade do equipamento. E por fim, a multiplicação do conhecimento, para que as equipes possam aplicar e melhorar os conceitos e técnicas de ciências de dados no dia a dia da produção.

Referências

- ARAÚJO, C. G. et al. **Investigation of a PID control for automatic compensation of angular deviation in a continuous annealing line.** *Metallurgical and Materials Transactions A*, v. 48, n. 3, p. 515-520, 2017.
- BERKSON, J. **Are there 'Pure' Chance Experiments?** *The Annals of Mathematical Statistics*, v. 15, n. 3, p. 205-218, 1944.
- BROWN, A.; MILLER, B. **Real-time monitoring systems for cold rolling oil quality.** *Journal of Manufacturing Science and Engineering*, v. 141, n. 5, p. 051004, 2019.
- BROWN, C.; WHITE, D. **Impact of Oil Contamination on Rolling Mill Performance.** In: *International Conference on Metalworking Technologies, Proceedings*, p. 123-135, 2021.
- BULGIN, M. N. **A brief history of the Sendzimir rolling mill.** ASM International, 1988.
- CHOLLET, F. **Deep Learning with Python.** Manning Publications, 2017.
- COX, D. R. **The regression analysis of binary sequences.** *Journal of the Royal Statistical Society. Series B (Methodological)*, v. 20, n. 2, p. 215-242, 1958.
- GARCIA, C. et al. **Recycling and regeneration of cold rolling oil: A sustainable approach.** *Resources, Conservation and Recycling*, v. 174, p. 105839, 2021.
- GARCIA, E.; MARTINEZ, F. **Optimization of Filtration Systems for Hot Rolling Processes.** *Journal of Manufacturing Engineering*, v. 25, n. 4, p. 78-92, 2018.
- GAUTHAM, P. R. **Metallography: Microstructural Analysis and Control.** McGraw-Hill, 2004.
- HOSMER, D. W.; LEMESHOW, S.; STURDIVANT, R. X. **Applied logistic regression.** John Wiley & Sons, 2013.
- IYER, N. V.; SMITH, C. C. **Advances in aluminum rolling: Production and end uses of advanced high-performance alloys.** ASM International, 2005.
- JADHAV, D. D. **A comprehensive guide to material science.** Spon Press, 2013.

JONES, R. **Centrifugal filtration for cold rolling oil purification.** *Chemical Engineering Research and Design*, v. 160, p. 152-160, 2020.

NANJIANG, J. **Handbook of High-Temperature Rolling: Mechanisms and Processes.** Springer, 2014.

OLUWOLE, J. A. **Fundamentals of Rolling.** Woodhead Publishing, 2006.

PEDREGOSA, F. et al. **Scikit-learn: Machine learning in Python.** *Journal of Machine Learning Research*, v. 12, p. 2825-2830, 2011.

SMITH, J.; JONES, A. **Advances in Rolling Oil Filtration Techniques.** *Journal of Steel Industry*, v. 10, n. 2, p. 45-60, 2021.

SMITH, J. et al. **High-efficiency cartridge filtration for cold rolling oil.** *Journal of Materials Processing Technology*, v. 252, p. 452-461, 2018.

TANENBAUM, A. S.; WETHERALL, D. J. **Redes de Computadores.** Elsevier, 2011.

TOKYO TECH. **Dual Degree Program.** 2005.

MORETTIN, P. A.; BUSSAB, W. O. **Estatística Básica.** Saraiva, 2017.