

FELIPE PEREIRA SOUSA

**USO DE MODELOS PREDITIVOS NÃO LINEARES NA
PREVISÃO DE SATURAÇÃO DE FILTROS *DIASTAR* EM
UMA REFINARIA DE ALUMINA**

São Luís
2026

FELIPE PEREIRA SOUSA

**USO DE MODELOS PREDITIVOS NÃO LINEARES NA
PREVISÃO DE SATURAÇÃO DE FILTROS *DIASTAR* EM
UMA REFINARIA DE ALUMINA**

Trabalho de Conclusão de Curso apresentado ao Colegiado de Curso da Engenharia Química do Centro de Ciências Exatas e Tecnologia da Universidade Federal do Maranhão, como parte dos requisitos para obtenção do diploma de Graduação em Engenharia Química.

Orientador: Prof. Dr. Romildo Martins Sampaio

São Luís
2026

Sousa, Felipe.
USO DE MODELOS PREDITIVOS NÃO LINEARES NA PREVISÃO DE
SATURAÇÃO DE FILTROS DIASTAR EM UMA REFINARIA DE ALUMINA /
Felipe Sousa. - 2026.
50 f.

Orientador(a): Romildo Sampaio.
Curso de Engenharia Química, Universidade Federal do
Maranhão, Auditório da Proen, 2026.

1. Processo Bayer. 2. Filtro Diastar. 3. Machine
Learning. 4. Multilayer Perceptron. 5. Random Forest. I.
Sampaio, Romildo. II. Título.

BANCA EXAMINADORA:

Prof. Dr. ROMILDO MARTINS SAMPAIO
Orientador – CCEQ/CCET/UFMA

Prof. Dr. HARVEY ALEXANDER VILA-VELLEZ
CCEQ/CCET/UFMA

Prof. Dr. PAULO HENRIQUE SILVA LEITE COELHO
CCEQ/CCET/UFMA

23 de janeiro de 2026

DADOS CURRICULARES

Felipe Pereira Sousa

NASCIMENTO 28/12/1999 – SÃO LUIS / MA

FILIAÇÃO Welinton de Oliveira Souza
Soraia Cunha Pereira Souza

2018/2025 Curso de Graduação
Engenharia Química - Universidade Federal do Maranhão

Dedico este trabalho à minha família, que sempre me apoiou no decorrer desta caminhada.

AGRADECIMENTOS

Aos meus pais, por me fornecerem todas as condições possíveis de focar nos meus estudos.

Aos meus amigos da escola D. Pedro II, por me acompanharem ao longo dessas décadas de amizade, por todos os altos e baixos que passei e sempre estarem lá por mim.

Aos amigos que fiz na UFMA no decorrer desses anos, com todo o suporte e apoio mútuo que construímos em todo esse processo de graduação.

Ao meu orientador, Prof. Romildo, pelo incentivo e confiança no meu trabalho.

Ao Prof. Harvey, pelos ensinamentos e pela disposição de sempre ajudar.

“Computadores são inúteis, eles só sabem dar respostas”

Pablo Picasso

SOUSA, Felipe P. **Uso de modelos preditivos não lineares na previsão de saturação de filtros diastar em uma refinaria de alumina**. 2026. 50 f. Trabalho de Conclusão de Curso de Engenharia Química do Centro de Ciências Exatas e Tecnologia da Universidade Federal do Maranhão, São Luís, 2026.

RESUMO

A produção de alumina pelo processo Bayer é limitada pela instabilidade das soluções supersaturadas utilizadas pelas refinarias de alumina. O processo de filtração de licor rico em alumina é vital para a qualidade do produto final, assim como a etapa mais impactada pelos níveis de supersaturação do licor. A elaboração de um modelo preditivo de saturação de filtros surgiu com a necessidade de um controle mais fino de turbidez e operação do equipamento em seu ponto ótimo. Foram desenvolvidos e comparados modelos de aprendizado de máquina, especificamente Multilayer Perceptron (MLP) e Random Forest, utilizando a linguagem Python. A base de dados utilizada compõe-se de conjuntos de dados históricos de filtros de uma refinaria, sendo a performance dos modelos avaliada através do coeficiente de determinação (R^2) e do erro médio absoluto. Os resultados indicaram que o modelo MLP apresentou desempenho superior ao Random Forest, obtendo um coeficiente de determinação (R^2) de 0,95 contra 0,86 do comparativo. As análises dos diferentes conjuntos no decorrer dos anos de 2022 a 2024 mostrou que a utilização de dados mais recentes no treinamento dos modelos é capaz de adequar predições à ajustes operacionais comuns no ambiente industrial.

Palavras-chave: Processo Bayer. Filtro Diastar. Machine Learning. Multilayer Perceptron. Random Forest.

SOUSA, F. P. Use of Nonlinear Predictive Models in Forecasting Diastar Filter Saturation in an Alumina Refinery. 2026. 50 f. Graduate Work (Graduate in Chemical Engineering) – Curso de Engenharia do Centro de Ciências Exatas e Tecnologia da Universidade Federal do Maranhão, São Luís, 2026.

ABSTRACT

The production of alumina by the Bayer process is limited by the instability of the supersaturated solutions used in alumina refineries. The filtration process of alumina-rich liquor is vital to the quality of the final product, as well as being the most affected stage by the liquor's supersaturation levels. The development of a predictive model for filter saturation arose from the need for finer control of turbidity and for operating the equipment at its optimal point. Machine learning models were developed and compared, specifically Multilayer Perceptron (MLP) and Random Forest, using the Python programming language. The dataset used consists of historical data from filters in a refinery, and model performance was evaluated using the coefficient of determination (R^2) and the mean absolute error. The results indicated that the MLP model outperformed the Random Forest, achieving a coefficient of determination (R^2) of 0.95 compared to 0.86 for the latter. Analyses of different datasets over the years 2022 to 2024 showed that using more recent data for model training enables predictions to better adapt to common operational adjustments in the industrial environment.

Keywords: Bayer Process. Diastar Filter. Machine Learning. Multilayer Perceptron. Random Forest.

LISTA DE FIGURAS

Figura 1 – Fluxograma do processo Bayer	17
Figura 2 – Filtro Kelly com novos panos instalados	19
Figura 3 – Esquema 3D de um filtro Diastar	23
Figura 4 – Esquema de como funciona o Random Forest	25
Figura 5 – Exemplo de como funciona uma MLP	27
Figura 6 – Funções de ativação mais utilizadas em redes neurais	28
Figura 7 – Métricas de desempenho por modelo e conjunto	35
Figura 8 – Matriz da correção de Pearson (C1)	36
Figura 9 – Matriz da correção de Pearson (C4)	37
Figura 10 – Matriz da correção de Spearman (C1)	38
Figura 11 – Matriz da correção de Spearman (C4)	39
Figura 12 – Fluxo do filtro previsto versus real com MLP no C1	40
Figura 13 – Fluxo do filtro previsto versus real com RF no C1	40
Figura 14 – Fluxo do filtro previsto versus real com MLP no C4	41
Figura 15 – Fluxo do filtro previsto versus real com RF no C4	42
Figura 16 – Previsão da taxa de filtração com MLP no C1	43
Figura 17 – Previsão da taxa de filtração para filtros 31, 35 e 39	44
Figura 18 – Previsão da taxa de filtração para os filtros 33, 37 e o total da bancada	45
Figura 19 – Previsão da taxa de filtração para um de filtro com alarme de saturação	46

LISTA DE TABELAS

Tabela 1 – Variáveis utilizadas.....	31
Tabela 2 – Descrição dos conjuntos de dados utilizados.....	32
Tabela 3 – Resultados obtidos em C1	46

LISTA DE ABREVIATURAS E SIGLAS

ICSOBA	International Committee for Study of Bauxite, Alumina and Aluminium
RF	Random Forest
FFNN	Feed Forward Neural Network
ReLu	Rectified Linear Unit
MLP	Multilayer Perceptron

LISTA DE SÍMBOLOS

x	variáveis do processo
k	neurônio
z	saída linear do neurônio
w	peso
b	viés
μ	média
σ	desvio padrão da variável

SUMÁRIO

FOLHA DE ROSTO.....	i
FICHA CATALOGRÁFICA.....	ii
FOLHA DE APROVAÇÃO.....	iii
DADOS CURRICULARES.....	iv
DEDICATÓRIA.....	v
AGRADECIMENTOS.....	vi
EPÍGRAFE.....	vii
RESUMO.....	viii
ABSTRACT.....	ix
LISTA DE FIGURAS.....	x
LISTA DE TABELAS.....	xi
LISTA DE ABREVIATURAS E SIGLAS.....	xii
LISTA DE SÍMBOLOS.....	xiii
SUMÁRIO.....	xiv
1 INTRODUÇÃO.....	16
1.1 Processo Bayer.....	16
1.2 Clarificação do licor.....	17
1.3 Filtragem do licor.....	18
1.4 Limitações da modelagem em reatores industriais.....	20
2 OBJETIVOS.....	21
2.1 Objetivos gerais.....	21
2.2 Objetivos específicos.....	21
3 REVISÃO BIBLIOGRÁFICA.....	22
3.1 Filtro Diastar.....	22
3.2 Modelagem preditiva: random forest e redes neurais artificiais.....	24
3.2.1 Random Forest (RF).....	24
3.2.2 Rede Neural Feed Forward (FFNN).....	25
3.3 Trabalhos Correlatos.....	29
4 METODOLOGIA.....	31
4.1 Local da pesquisa.....	31
4.2 Conjunto de dados.....	31

4.3	Tratamento de dados.....	32
4.4	Análises das correlações.....	33
4.5	Técnicas de modelagem.....	33
4.6	Análises estatísticas.....	34
5	RESULTADOS E DISCUSSÃO.....	35
6	CONCLUSÃO.....	47
	REFERÊNCIAS.....	48

1 INTRODUÇÃO

1.1 PROCESSO BAYER

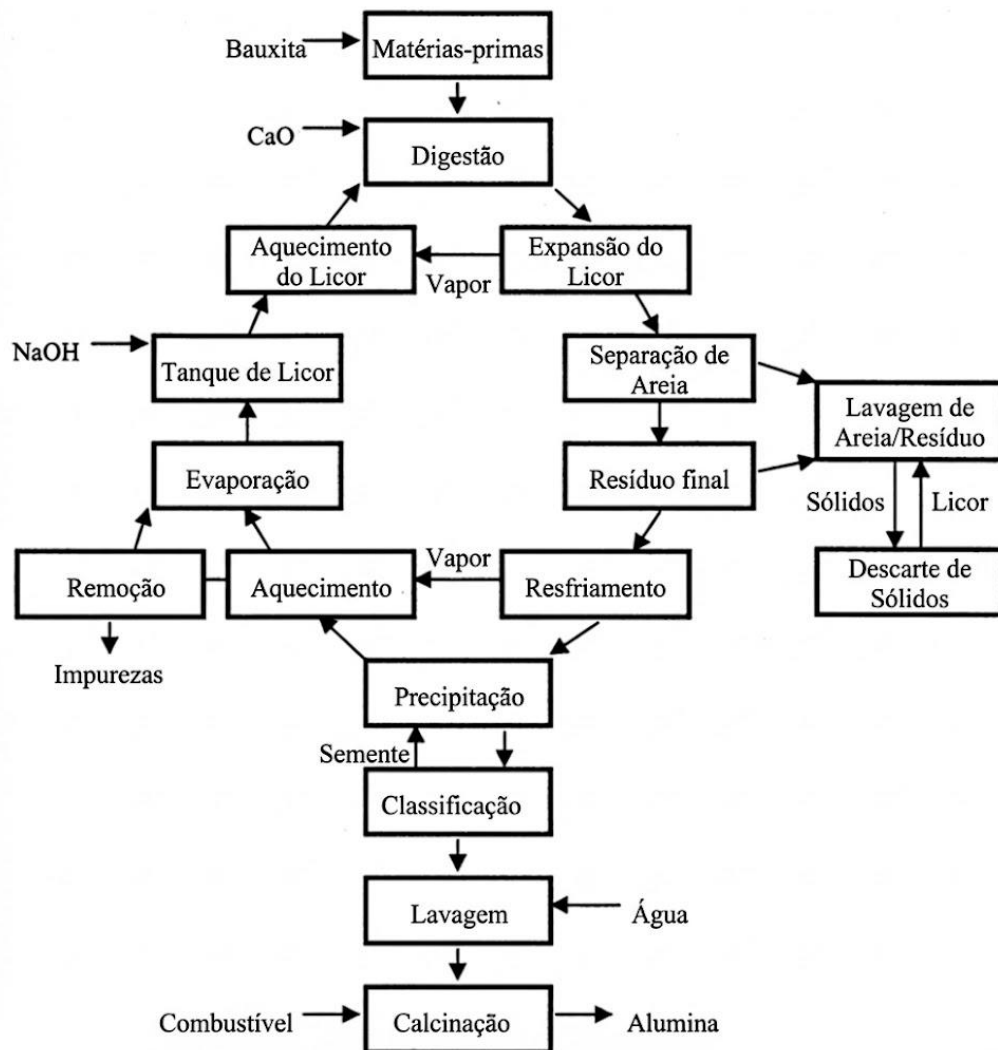
O processo Bayer, patenteado em 1888 pelo austríaco Karl Joseph Bayer, é utilizado no refino da bauxita e obtenção de alumina, empregando soda cáustica sob alta pressão e temperatura. A alumina é dissolvida na soda cáustica formando uma solução chamada de licor, que é enriquecido de alumina à medida que a digestão da bauxita ocorre, sendo precipitada no final do processo na forma de hidróxido de alumina (Habashi, 2013; Hind; Bhargava; Grocott, 1999).

A bauxita é um minério composto de diversos materiais que possui predominantemente o alumínio na sua composição. Dentre os componentes majoritários está o alumínio, o ferro, o titânio e a sílica. Existem diferentes tipos de bauxita tanto com relação à composição quanto à forma do mineral do alumínio, sendo elas a gibbsita, a bohemita e a diásporo. A forma mais valiosa e preferencial do minério é a gibbsita por requerer uma menor temperatura para sua extração, cerca de 150 °C (Hind; Bhargava; Grocott, 1999).

Uma refinaria comum de alumina baseia-se em uma série de processos que transformam essa matéria prima em produto final, como mostra a Figura 1. As etapas podem ser simplificadas em moagem, digestão, clarificação, precipitação e calcinação, acompanhados de processos de aquecimento ou resfriamento do licor a depender da etapa em que se encontra. O processo como um todo é cíclico tendo o aproveitamento do licor pobre em alumina obtido no final do processo de volta no início. Apesar da soda cáustica ser reaproveitada, ela é constantemente consumida pela sílica presente no minério, sendo necessário a dosagem de soda para que ocorra novamente o processo de extração (Hind; Bhargava; Grocott, 1999).

A parte do minério que não é aproveitada é chamada de lama vermelha que é eliminada no processo de clarificação do licor. A quantidade de resíduo é diretamente dependente da quantidade de alumínio presente no minério, ou seja, quanto menor a porcentagem de alumínio, maior será a taxa de resíduo. Este resíduo é depositado em áreas próprias para o descarte conhecidas como lagoa de resíduos, que são construídas para suportar um determinado volume desse material dentro do seu limite estrutural. O descarte corresponde a uma parcela significativa do custo de produção de alumina, custo esse relacionado não somente à construção e operação da área, mas também ao valor da soda perdida ligada ao resíduo e o custo de neutralizar e reabilitar um ambiente hostil após a conclusão do descarte (Cooling, 2016).

Figura 1: Fluxograma do processo Bayer.



Fonte: Adaptado de Hind (1999, p. 362)

1.2 CLARIFICAÇÃO DO LICOR

A etapa da clarificação é composta por espessadores, lavadores e filtros, que são responsáveis por separar os sólidos do licor, além de recuperar alumina e soda presentes no rejeito antes de chegarem à lagoa de resíduo. Esta etapa recebe da digestão um licor quente e rico em alumina em conjunto com o resíduo. Ao se tratar de processos de decantação e filtração, quaisquer variações bruscas de temperatura, fluxo ou qualidade da bauxita, acarretam em perda de floculação (Hind; Bhargava; Grocott, 1999).

Os espessadores são a primeira etapa desse processo e realizam a decantação do rejeito com auxílio de floculantes sintéticos, tendo sua corrente de topo um licor com baixo teor de sólidos para os filtros e na corrente de fundo a lama vermelha para a cadeira de lavadores. A

operação deste tanque em suas condições ótimas é essencial para as próximas etapas da produção (Patil; Thorat, 2022).

Os lavadores recebem, além da lama vermelha, água proveniente da lagoa de resíduos. Esta água contém um certo teor de alumina e soda que são reinsertos no processo. O processo de lavagem da lama vermelha ocorre em contracorrente, com a lama seguindo na direção do lago e a água seguindo na direção dos espessadores. Apesar da lavagem, nem toda alumina e soda são recuperadas nesse processo, possuindo assim uma porcentagem fixa de perda para o resíduo (Patil; Thorat, 2022).

A última etapa que separa o licor do rejeito é a filtração, que por sua vez é crucial para a qualidade do produto final, uma vez que impurezas presentes na alumina afetam a qualidade do alumínio produzido. A filtração age como a última barreira de impurezas presentes no licor, assim como produtos empregados no processo produtivo, como os flocculantes sintéticos discutido anteriormente e o cálcio, que é utilizado como um agente estabilizador do licor e um auxiliar de formação da torta no processo de filtração (Bott; Langeloh; Hahn, 2016).

1.3 FILTRAGEM DO LICOR

A filtração de licor rico em alumina é um processo caracterizado por vazões altas, uma concentração de sólidos em média menor que 300 mg/L, temperaturas acima de 100 °C e pressões absolutas acima de 5 bar. Os tipos de filtros mais comuns presentes nas refinarias são os filtros Kelly ou filtros de pressão horizontal, como mostrado na Figura 2, que operam em bateladas de 5 a 10 horas que representam o tempo de saturação do pano que realiza a filtração. Para as demandas atuais do setor de alumina, tem-se o *Diastar* como sendo o mais importante na filtração de licor, por ser capaz de suportar altas temperaturas e altas taxas de fluxo. Os panos possuem uma certa porosidade que impede a passagem de partículas maiores, que por sua vez formam uma torta que captura partículas menores, ou seja, tem-se a torta como o principal elemento filtrante desse tipo de equipamento (Bott; Langeloh; Hahn, 2016).

Os ciclos envolvem a formação da torta no pano ao longo da filtração com o aumento da resistência da torta crescente ao decorrer do processo, o que se denomina de saturação do filtro. O tempo de saturação depende do tipo de filtro e das características do licor. Quanto mais próximo da supersaturação do licor, mais ele estará mais propenso à autoprecipitar alumina durante o processo, categorizando perda de produção, além de causar rasgos nos panos devido à característica da alumina em formar cristais ao precipitar. Atingir estados de supersaturação requer um controle fino de dosagem de soda, bauxita e temperatura (Bach;

Schoenbrunn, 2022; LaMacchia; Costa, 2022).

Figura 2: Filtro Kelly com novos panos instalados.



Fonte: MarketGroup, (2025)

Com o tempo de operação ocorre a saturação do meio filtrante por sólidos que entopem os poros ao longo do tempo ou por floculantes que foram arrastados dos espessadores por excesso de dosagem ou temperatura (fervura), grudando no pano e cegando a área atingida. A saturação do pano resulta em aumento de resistência pelo próprio pano e em redução do fluxo que o equipamento é capaz de filtrar, acarretando em rasgos em situações extremas. Rasgos podem ser identificados por aumento súbito da vazão ou turbidez do licor na saída (Gontijo *et al.*, 2009).

Os panos são o material mais caro da manutenção deste tipo de equipamento, com custo aproximado de R\$ 160 mil, somente de material, em cada troca. Atualmente, na planta onde este trabalho foi desenvolvido, a troca é feita de forma corretiva ou após 70 dias de operação. Com um número total de 10 filtros e trocas individuais feitas a cada 70 dias, tem-se uma troca por semana ao decorrer do ano de produção, sendo possível prolongar a vida útil quando o filtro passa por lavagens e ciclos corretos sem extrapolar seus limites de resistência. Para cada dois dias adicionais de vida útil dos panos, há uma redução no número de trocas e paradas de manutenção de duas semanas por ano, o que representa um ganho de aproximadamente R\$ 320 mil somente de material. Sendo assim a elaboração de um modelo preditivo que permite dizer quando se deve retirar cada filtro de operação é vital para otimizar

custos, reduzir cortes de produção e evitar paradas não planejadas.

1.4 LIMITAÇÕES DA MODELAGEM EM REATORES INDUSTRIAIS

A modelagem de reatores químicos industriais tem sido tradicionalmente fundamentada em modelos derivados de balanços de massa e energia e leis fundamentais da termodinâmica e cinética. No entanto, a aplicação desses modelos em cenários industriais reais enfrenta obstáculos significativos. Reatores industriais são sistemas complexos, caracterizados por dinâmicas fortemente não lineares, interações multivariáveis, cinéticas de reação mal compreendidas e variações nos fluxos, temperaturas ou pressões que regem o processo. Nessas condições, a obtenção de modelos precisos torna-se uma tarefa desafiadora e, em muitos casos, economicamente inviável (Butean *et al.*, 2025).

O avanço das técnicas de Inteligência Artificial (IA) tem possibilitado o desenvolvimento de modelos capazes de aprender padrões complexos a partir de dados, ampliando significativamente o potencial de sistemas preditivos em processos industriais e científicos. Nesse contexto, métodos baseados em Aprendizado de Máquina, especialmente aqueles fundamentados em Redes Neurais Artificiais, têm se destacado por sua capacidade de modelar relações não lineares entre variáveis. Diante desse cenário, este trabalho propõe o desenvolvimento de um modelo preditivo utilizando tais técnicas, visando a predição da saturação dos filtros *Diastar* (Li *et al.*, 2025).

2 OBJETIVOS

O presente trabalho busca desenvolver um modelo capaz de prever o comportamento de filtros *Diastar*, através de técnicas de inteligência artificial, mais especificamente métodos de aprendizado de máquina baseados em redes neurais, a fim de identificar e operar o equipamento em seu ponto ótimo, permitindo uma estratégia operacional eficiente que prolonga sua vida útil, reduzindo custos e perdas de produção, e aprimorando o planejamento de manutenção preventiva.

2.1 OBJETIVO GERAL

Desenvolver um modelo capaz de prever o comportamento de um filtro *Diastar* em uma refinaria de alumina.

2.2 OBJETIVOS ESPECÍFICOS

- Analisar a relação das diferentes variáveis do processo na predição do filtro;
- Coletar e tratar dados operacionais necessários para o desenvolvimento do modelo;
- Ajustar e validar diferentes abordagens de modelagem;
- Avaliar o desempenho do modelo através de métricas estatísticas.

3 REVISÃO BIBLIOGRÁFICA

3.1 FILTRO *DIASTAR*

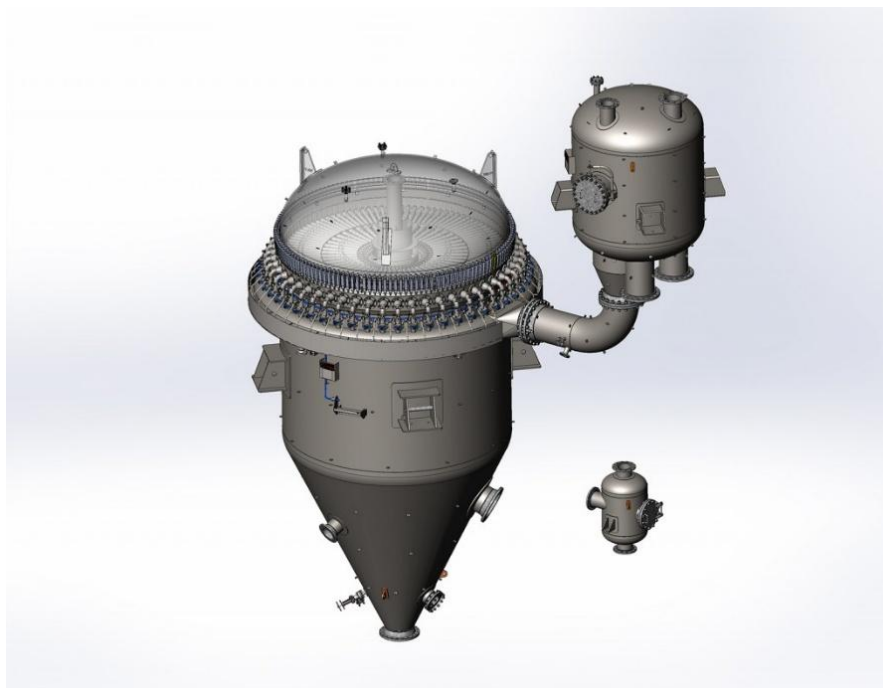
A produção de alumina no mundo é limitada a duas reações principais, a precipitação de gibsite e a cristalização de oxalato de sódio. A depender do tipo de bauxita empregado na produção é que se consegue estabelecer o quão concentrado será o licor no processo. Na Austrália, por exemplo, utiliza-se uma bauxita pobre em alumina e rica em matéria orgânica, necessitando etapas extras para a remoção dessas impurezas, tendo sua produção limitada pela capacidade de lidar com um volume maior de rejeito. Já no Brasil, tem-se uma bauxita mais pura, sem a presença de oxalatos e porcentagens de alumina que beiram os 50%. Nesse caso a produção é limitada apenas pela supersaturação do licor (Hind; Bhargava; Grocott, 1999; Komlóssy; Van Deursen; Raahauge, 2022).

A concentração do licor é a principal preocupação da etapa da filtração devido à possibilidade de formar casca e precipitar alumina nos panos. A casca é formada majoritariamente de rejeitos insolúveis formados no processo de digestão da bauxita, que se agarram nas paredes das tubulações, tanques e também nos panos dos filtros, atingindo com maior frequência a chamada área vermelha da refinaria, que lida justamente com o rejeito. Essa formação prejudica taxas de transferência de calor e restringem o fluxo nas áreas afetadas. Quanto mais concentrado for o licor, maior será a produção da indústria, porém consequentemente maiores serão os danos causados pelas incrustações, sendo necessário a atuação da indústria em busca de meios que reduzam sua ocorrência ao máximo (Cheng *et al.*, 2021).

O filtro *Diastar*, produzido pela empresa Gaudfrin, é um equipamento vertical com folhas com pressão estática, totalmente automático. Pode ser utilizado para a clarificação de suspensões ou lavagem e desidratação de tortas de filtração. Caracterizado por uma série de panos ao longo da sua circunferência, que retêm o material a ser filtrado, e um *headtank* (tanque que armazena uma certa quantidade de licor, conforme mostrado na Figura 3), que recebe o licor clarificado. Seu funcionamento consiste em filtrar por ciclos de uma hora seguidos por um desprendimento da torta pela pressão hidrostática do líquido armazenado no *headtank*. Esse padrão segue até atingir um ponto de saturação onde a torta já não desprende mais com a força do líquido, sendo necessário a lavagem do filtro para entrar em funcionamento novamente. A lavagem é feita com solução cáustica para digerir a alumina no resíduo e preparar o filtro para entrar em funcionamento novamente. Resíduos separados nessa etapa, em sua maioria, são reinsertos na clarificação para recuperar parte da alumina e soda ainda presentes (Bach;

Schoenbrunn, 2022; LaMacchia; Costa, 2022).

Figura 3: Esquema 3D de um filtro *Diastar*.



Fonte: Gaudfrin (2025).

Os *Diastar* são considerados o estado da arte quando se trata da filtração de licor por: permitir um menor tempo de residência na filtração; permitir um acúmulo da camada de lama desalojada do pano durante ciclos com descarga sem necessidade da adição de água; operar em condições extremas de temperatura e pressão; e reduzir o tempo fora de operação por filtrar e lavar de maneira automática. Ao invés de operarem em ciclos de 5 a 10 horas, como no filtro *Kelly*, o *Diastar* pode operar com tempos de filtração inferiores a 1 hora de forma contínua. Devido a esses fatores, somente esse filtro permite operar uma refinaria de alumina em condições extremas de supersaturação do licor (LaMacchia; Costa, 2022).

Melo *et al.* (2023), em estudo apresentado na 41ª Conferência ICSOBA, investigou a modernização das etapas de limpeza e separação de resíduos na produção de alumina, utilizando como base as operações da refinaria Hydro Alunorte. A pesquisa compara o método tradicional, que depende de grandes tanques onde as impurezas se depositam no fundo por ação da gravidade, com uma tecnologia mais recente de filtros que funcionam sob pressão vertical (*Diastar*). Os autores concluíram que o uso destes filtros modernos oferece vantagens operacionais significativas sobre o sistema antigo, resultando num líquido final muito mais limpo e num melhor aproveitamento dos produtos químicos do processo, o que demonstra ser uma solução eficiente para indústrias que procuram aumentar a sua produtividade e reduzir

desperdícios sem a necessidade de expandir fisicamente as suas áreas de tanques.

Outros autores consideram o filtro *Kelly* como o estado da arte pelo fato de ser o mais comum e mais utilizado ao longo dos anos por refinarias ao redor do mundo. Porém, esta tecnologia deixa a desejar em termos de desempenho e operação do filtro, como em longos períodos de inatividade, qualidade da alumina, grande necessidade de mão de obra (lavagens manuais) e baixa performance do filtro. Por isso, algumas refinarias, como a de Clarendon na Jamaica, substituem seus filtros *Kelly* ou filtros de areia, por filtros *Diastar* (Bach; Schoenbrunn, 2022; Bott; Langeloh; Hahn, 2016; Shaw; Duncan; Crosdale, 2011).

3.2 MODELAGEM PREDITIVA: *RANDOM FOREST* E REDES NEURAIAS ARTIFICIAIS

A modelagem de fenômenos complexos frequentemente expõe as limitações das abordagens estatísticas clássicas, que, ao assumirem linearidade, podem falhar em capturar a estrutura real dos dados. Nesse contexto, Breiman (2001) argumenta que a adoção de uma modelagem algorítmica — em contraposição à modelagem de dados tradicional — é essencial para lidar com sistemas de alta dimensionalidade, pois permite identificar padrões não-lineares intrincados sem a imposição de formas funcionais rígidas. Essa capacidade de generalização e foco na precisão preditiva fundamenta a escolha de métodos robustos como o Random Forest e o Multilayer Perceptron, detalhados a seguir (Breiman, 2001).

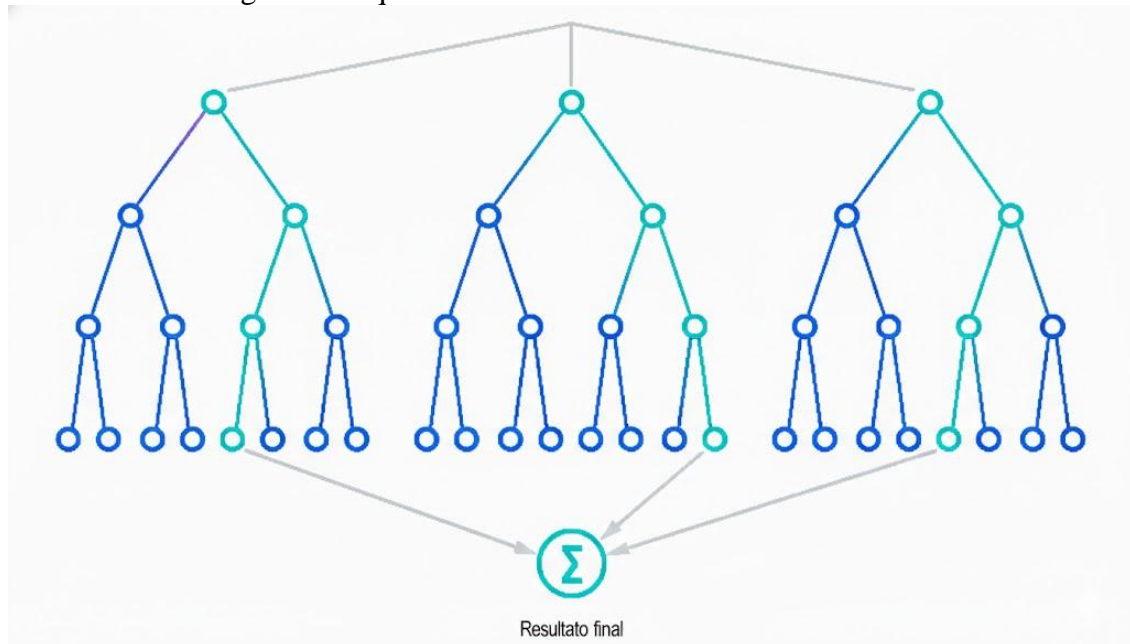
3.2.1 *Random Forest* (RF)

O *Random Forest* é um modelo baseado em aprendizado de máquina supervisionado que combina múltiplas árvores de decisão, em uma técnica chamada *ensemble* (conjunto), para gerar previsões mais robustas, como mostrado na Figura 4. Sua principal vantagem reside na capacidade de modelar relações complexas e não lineares entre as variáveis, mesmo na presença de ruído e alta variabilidade nos dados. Essa flexibilidade permite que o *Random Forest* apresente desempenho superior em cenários onde modelos mais simples falham, resultando em previsões mais precisas e estáveis (Breiman, 2001; Schonlau; Zou, 2020).

A estrutura fundamental de uma única árvore opera particionando o espaço de dados em regiões retangulares, iniciando em um nó raiz, onde se encontra a totalidade do conjunto de dados, passando por nós de decisão, onde ocorrem as divisões baseadas em atributos, até alcançar os nós folha, que representam o resultado final da predição (classe ou valor numérico). A estratégia de utilizar um *ensemble* visa mitigar as limitações individuais de uma única árvore, resultando em um modelo global mais robusto e com maior capacidade de generalização para

problemas complexos de engenharia (Sarker, 2021).

Figura 4: Esquema de como funciona o *Random Forest*.



Fonte: Adaptado de IBM (2025)

A consolidação dos resultados individuais de cada árvore para gerar a previsão final do modelo ocorre através de métodos de agregação, que variam conforme a natureza do problema (classificação ou regressão). Para tarefas de regressão, o resultado final é obtido pela média aritmética das previsões de todas as árvores da floresta. Já em tarefas de classificação, o modelo opera por votação majoritária, onde a mais frequente entre as árvores é eleita como a saída final. Essa abordagem democrática assegura que erros pontuais de árvores individuais sejam compensados pelo acerto coletivo da floresta, resultando em previsões de alta acurácia e estabilidade frente a *outliers* (valores atípicos) (Breiman, 2001).

Embora o *Random Forest* seja composto por árvores de decisão individualmente interpretáveis, o grande número de modelos e o processo de agregação das previsões fazem com que o modelo final apresente características de *black box*, dificultando a interpretação direta da relação entre variáveis de entrada e saída. Tendo como principais pontos fracos a dependência de dados de boa qualidade e risco de *overfitting* (sobreajuste), se ajustando demais aos dados do treinamento e performando mal em novos dados (Dobbelaere *et al.*, 2021).

3.2.2 Rede Neural Feed Forward (FFNN)

A Rede Neural *Feed Forward* (FFNN) é um modelo computacional composto por camadas sequenciais de unidades de processamento, chamadas de neurônios, que realizam

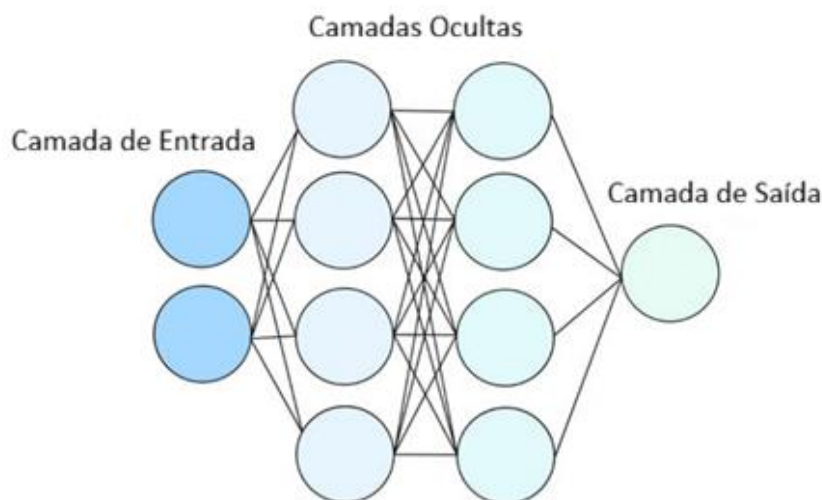
transformações sobre os dados de entrada até gerar a saída desejada. Uma representação visual do modelo é mostrada na Figura 5. Esse tipo de rede é especialmente indicado para problemas que envolvem relações complexas e não lineares entre as variáveis, nas quais métodos tradicionais apresentam limitações. Durante o treinamento, os pesos das conexões entre os neurônios são ajustados automaticamente, com o objetivo de minimizar o erro nas previsões. Trata-se de um modelo altamente versátil, que pode ser adaptado conforme a complexidade do problema por meio do ajuste do número de camadas ocultas e neurônios, permitindo maior capacidade de representação e generalização dos dados (Dutta *et al.*, 2018).

O *Multilayer Perceptron* (MLP) é um tipo de FFNN que opera em múltiplas camadas totalmente conectadas. Cada neurônio em um MLP opera realizando uma soma ponderada de suas entradas, acrescida de um termo chamado *bias* (viés), seguida pela aplicação de uma função matemática. Formalmente, para um dado neurônio \mathbf{k} , se as entradas são representadas pelo vetor \mathbf{X} os pesos sinápticos por \mathbf{w} e o viés por \mathbf{b} , o processamento linear inicial pode ser descrito pela equação (1). Os pesos (w) são os parâmetros ajustáveis que a rede "aprende" durante o treinamento, determinando a intensidade da influência de cada variável de entrada na saída do neurônio (Rynkiewicz, 2012).

$$z_k = \sum(w_i \cdot x_i) + b \quad (1)$$

A estrutura que caracteriza o MLP como uma rede "multicamadas" organiza esses neurônios em uma topologia hierárquica composta por: uma camada de entrada (que recebe os dados brutos dos sensores), uma ou mais camadas ocultas (*hidden layers*) e uma camada de saída. LeCun; Bengio; Hinton, (2015) explicam que as camadas ocultas são o motor principal do modelo, responsáveis por distorcer a entrada de forma não-linear para tornar as classes ou valores linearmente separáveis na saída. Enquanto redes sem camadas ocultas são limitadas a resolver problemas lineares, a adição dessas camadas permite que o MLP atue como um aproximador universal de funções, capaz de capturar relações complexas e interações não-lineares típicas de processos químicos industriais.

Figura 5: Exemplo de como funciona uma MLP



Rede neural feedforward totalmente conectada

Fonte: Adaptado de DataCamp (2025).

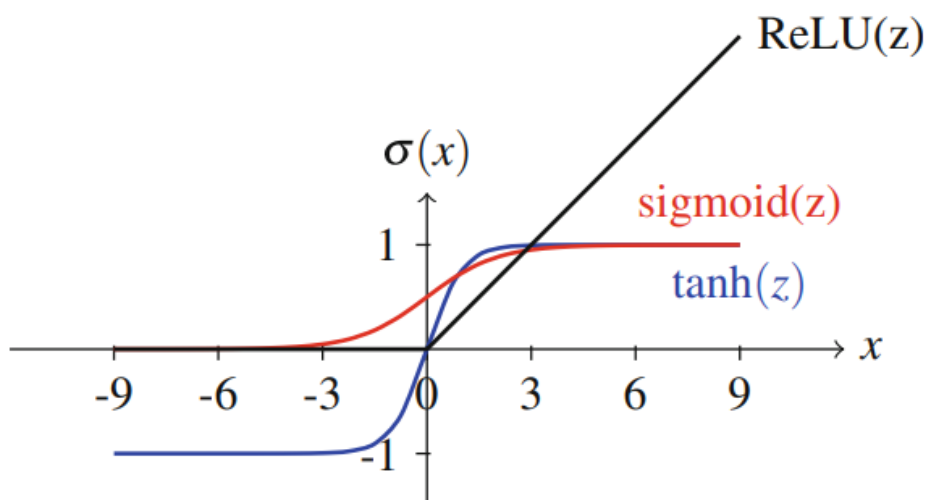
Para que a rede possa modelar essas não-linearidades, é imperativo o uso de funções de ativação aplicadas à saída de cada neurônio. Sem elas, independentemente da profundidade da rede, o modelo se comportaria como uma simples regressão linear. Entre as principais funções tem-se a *Rectified Linear Unit* (ReLU), a Sigmoid e a Tanh, como mostrado nas equações (2), (3) e (4). A ReLU retorna o próprio valor de entrada se ele for positivo; caso contrário, retorna zero. A Sigmoid é uma curva em forma de "S" que esmaga qualquer valor de entrada para o intervalo entre 0 e 1, sendo ideal para representar probabilidades. Já a Tanh é similar à Sigmoid, mas ajusta os valores para o intervalo entre -1 e 1, o que centraliza os dados no zero e geralmente acelera o aprendizado da rede. Para a camada de saída, a escolha da função depende da natureza do problema. Em tarefas de regressão, comumente utiliza-se a função Linear, permitindo que a rede preveja valores contínuos reais, enquanto a função Sigmoid é frequentemente aplicada em classificações binárias. A Figura 5 mostra a diferença gráfica de cada função de ativação citada (Dutta *et al.*, 2018; Szandala, 2021).

$$\text{ReLU}(x) = \max(0, x) \quad (2)$$

$$\text{Sigmoid}(x) = \frac{1}{1+e^{-x}} \quad (3)$$

$$\text{Tanh}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (4)$$

Figura 6: Funções de ativação mais utilizadas em redes neurais.



Fonte: Dutta *et al.*, (2018).

O aprendizado da rede ocorre efetivamente através do algoritmo de *Backpropagation* (Retropropagação). Este processo compara a predição da rede com o valor real utilizando uma Função de Perda (*Loss Function*), como o Erro Quadrático Médio (MSE). O algoritmo então calcula o gradiente do erro em relação a cada peso da rede, propagando-o da saída de volta para a entrada. Para atualizar os pesos e minimizar o erro, utilizam-se otimizadores. Atualmente, o algoritmo Adam (*Adaptive Moment Estimation*) é amplamente recomendado na literatura por adaptar a taxa de aprendizado individualmente para cada parâmetro, convergindo mais rapidamente que o clássico Gradiente Descendente Estocástico (SGD) em problemas com grandes volumes de dados e ruído (Singarimbun; Nababan; Sitompul, 2019).

A normalização dos dados constitui uma etapa mandatória para o treinamento de modelos baseados em otimização de gradiente, como o MLP, embora seja dispensável para algoritmos baseados em árvores de decisão, como o *Random Forest*. No contexto das redes neurais, variáveis de entrada com ordens de grandeza díspares (ex: temperatura em graus vs. pressão em Pascal) podem distorcer a superfície de erro, fazendo com que o algoritmo de otimização priorize variáveis de maior magnitude e prejudique a convergência do *Backpropagation* ou sature as funções de ativação. Em contrapartida, o *Random Forest* é invariante à escala monotônica dos dados, pois suas divisões são baseadas em limiares de corte e não em distâncias euclidianas ou produtos escalares. Para adequar os dados ao MLP, utilizam-se frequentemente dois métodos: o Min-Max Scaling, que comprime os dados no intervalo [0, 1] preservando a distribuição original, ou a Padronização (Z-Score), que transforma os dados

para que tenham média zero e desvio padrão unitário, sendo este último mais robusto a *outliers*. As equações (5) e (6) mostram os métodos citados respectivamente (Singh; Singh, 2020).

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (5)$$

$$x' = \frac{x - \mu}{\sigma} \quad (6)$$

onde μ é a média e σ é o desvio padrão da variável.

3.3 TRABALHOS RELACIONADOS

A aplicação de técnicas de *Machine Learning* na indústria de processos químicos tem crescido exponencialmente, impulsionada pela necessidade de otimização em cenários onde a modelagem clássica encontra limitações. A literatura recente divide-se majoritariamente entre abordagens baseadas em árvores de decisão para diagnósticos interpretáveis e redes neurais para modelagem não-lineares.

No contexto da valorização energética de resíduos, Kamal Abdelbasset *et al.* (2022) investigaram a otimização da produção de biodiesel a partir de óleo vegetal residual usando um catalisador heterogêneo. O estudo aplicou e comparou diferentes técnicas computacionais, com destaque para o MLP e o RF, para modelar a influência de parâmetros críticos operacionais. Os resultados obtidos evidenciaram a eficácia das abordagens de inteligência artificial sobre os métodos estatísticos clássicos, tendo o modelo MLP demonstrado uma precisão excepcional com um coeficiente de determinação (R^2) de 0,998. Esta elevada capacidade preditiva permitiu mapear com rigor a superfície de resposta do processo, facilitando a identificação das condições ótimas para maximizar o rendimento da reação.

Para a produção de ácido fosfórico, Saaidi, Bichri e Abderafi (2023) utilizaram modelos de *machine learning* para predição da viscosidade dinâmica durante o processo industrial. Foram avaliados diferentes modelos computacionais, como redes neurais artificiais (ANN, similar ao MLP só que com apenas uma camada), *gradient boosting* e RF, utilizando dados reais do processo, considerando variáveis como teor de sólidos, taxa de cisalhamento e temperatura. Apesar de todos os modelos apresentarem bons resultados, o *gradiente boosting* foi o que apresentou o melhor desempenho preditivo. O estudo concluiu que a aplicação de técnicas de inteligência artificial é uma alternativa eficiente aos modelos tradicionais, contribuindo para a otimização e melhor compreensão do comportamento reológico no processo de produção de ácido fosfórico.

A análise dos trabalhos correlatos revela que ambos modelos são amplamente utilizados na indústria e capazes de atuar na predição de processos complexos. Poucos trabalhos, no entanto, comparam diretamente o desempenho de ambos os algoritmos na predição específica de saturação de filtros industriais. Este trabalho visa preencher essa lacuna, avaliando qual abordagem oferece o melhor compromisso entre precisão preditiva e complexidade computacional para o problema proposto.

4 METODOLOGIA

4.1 LOCAL DA PESQUISA

A pesquisa foi realizada em uma grande indústria do setor metalúrgico localizada em um polo industrial no estado do Maranhão. A fábrica opera 24 horas por dia e tem uma alta capacidade de produção de alumina, chegando à média de produção de mais de 10 mil toneladas por dia. Esse local foi escolhido porque possui os processos e equipamentos ideais para a aplicação prática deste estudo.

4.2 CONJUNTO DE DADOS

Os dados utilizados neste estudo estão organizados em 11 variáveis, conforme mostrado na Tabela 1.

Tabela 1: Variáveis utilizadas.

Identificação	Variável	Origem	Descrição
1	Fluxo do Filtro	Online	Vazão de entrada no Filtro
2	Tempo de Operação	Online	Tempo operando desde a última lavagem
3	Tempo de Campanha	Online	Tempo de utilização do filtro desde a última troca de panos
4	Tempo de Ciclo	Online	Tempo individual de cada ciclo dentro do tempo de operação
5	Filter-aid	Laboratório	Dosagem de CaO misturado com licor para auxílio na formação da torta
6	Sólidos	Laboratório	Concentração de sólidos no licor
7	Temperatura	Online	Temperatura do licor
8	Concentração A/TC	Laboratório	Concentração do licor em Alumina/Teor Caústico
9	Entrada de Licor na Digestão	Online	Vazão de licor na entrada da primeira etapa do processo bayer
10	Pressão	Online	Pressão medida dentro do filtro em kPa
11	Taxa de Queda do Headtank	Online	Vazão da descarga de licor presente no Backflush para os panos

Fonte: Próprio autor (2025).

As variáveis utilizadas tinham como origem instrumentos no campo ou análises de amostras coletadas pelo próprio laboratório da refinaria. O conjunto total das variáveis pode

ser separado entre aquelas intrínsecas ao licor, como as 5, 6, 7, 8 e 9. O restante são variáveis próprias do filtro. Dentre elas, a variável 1 é a principal do modelo e qual se quer obter a partir da predição.

As variáveis utilizadas no trabalho já foram pré-selecionadas dentre um grupo maior de variáveis atreladas a outras etapas do processo que antecede ou procede a filtração, que após não surtirem efeito na predição foram descartadas, sobrando apenas as mais relacionadas à própria filtração.

Os dados estão organizados em quatro conjuntos, cada um com características temporais específicas, conforme mostrado na Tabela 2. Os Conjuntos de dados 1 a 3 (C1 a C3) correspondem aos anos de 2022 a 2024, com amostragem de dados realizadas a cada 10 min ao longo de cada dia do ano. Cada conjunto representa um ano distinto dentro desse intervalo, permitindo uma análise detalhada do comportamento diário das variáveis monitoradas.

O Conjunto 4 reúne os anos de 2023 e 2024 em um único conjunto consolidado o que o torna especialmente útil para a identificação de tendências de longo prazo, além de apresentar diversas situações pelas quais a planta passou. Este conjunto não utilizou os dados do ano de 2022 devido à ausência do instrumento que mede a temperatura do licor na entrada dos filtros, sendo necessário, neste período específico, a utilização de um medidor mais distante do equipamento analisado, no caso em questão, o do espessador que antecede esta etapa do processo. Conforme apresentado na Tabela 2, os conjuntos analisados foram construídos utilizando as variáveis disponíveis no conjunto de dados.

Tabela 2: Descrição dos conjuntos de dados utilizados.

Conjunto	Ano	Intervalo	Variáveis
C1	2024	10 min	Todas
C2	2023	10 min	Todas
C3	2022	10 min	Todas com exceção da variável 7
C4	2023-2024	10 min	Todas

Fonte: Próprio Autor (2025).

Para a utilização desses modelos, tem-se uma divisão entre dados de treinamento e de previsão do modelo. Os conjuntos citados acima foram apenas para a parte de treinamento, sendo utilizado na modelagem os dados da planta das últimas 24 horas para a previsão.

4.3 TRATAMENTO DOS DADOS

Os dados brutos de treinamento do modelo foram coletados originalmente com

frequência de amostragem de 1 minuto. Para a etapa de análise, realizou-se uma reamostragem temporal (resampling) dos dados para intervalos regulares de 10 minutos. O método de agregação utilizado foi a média aritmética dos valores dentro de cada janela de tempo não sobreposta, atuando na redução da dimensionalidade dos dados e como um filtro passa-baixa, e reduzindo ruídos de alta frequência comuns na aquisição de dados minuto a minuto. Após a limpeza dos dados, utilizou-se o método Min-Max *scaling* para a normalização dos dados antes de inserir na rede neural.

4.4 ANÁLISE DE CORRELAÇÕES

A análise de correlação entre as variáveis do sistema foi uma etapa fundamental para identificar como os dados se relacionam entre si. Esse processo permitiu entender quais variáveis de entrada possuem maior influência na saturação dos filtros, auxiliando no direcionamento das análises e na seleção dos dados mais relevantes para os modelos.

Para isso, aplicou-se primeiramente o coeficiente de Pearson, que mede o grau de relação linear entre duas variáveis, com uma métrica que varia de -1 a 1, sendo 1 uma relação perfeita. Ele foi utilizado para identificar se, quando uma variável aumenta, a outra aumenta ou diminui em uma proporção constante. Essa métrica é eficaz para detectar relações diretas e simples nos dados da planta industrial (Thirumalai; Chandhini; Vaishnavi, 2017).

Complementarmente, utilizou-se a correlação de Spearman, que avalia relações baseadas em postos. Diferente de Pearson, o coeficiente de Spearman consegue identificar relações chamadas de monotônicas, ou seja, situações em que as variáveis caminham juntas, mas não necessariamente de forma linear ou proporcional. Isso foi essencial para captar comportamentos mais complexos nos sensores que o modelo de Pearson poderia ignorar (Thirumalai; Chandhini; Vaishnavi, 2017).

4.5 TÉCNICAS DE MODELAGEM

A modelagem foi desenvolvida utilizando a linguagem de programação Python, na versão 3.11. Para o desenvolvimento e execução dos scripts, utilizou-se um computador equipado com processador Intel Core i5, 8 GB de memória RAM e sistema operacional Windows 11.

Para a organização e o tratamento da base de dados, foram utilizadas as bibliotecas Pandas e NumPy. O Pandas permitiu estruturar os dados históricos em tabelas, facilitando a limpeza e a preparação das variáveis, enquanto o NumPy foi empregado para realizar as

operações matemáticas e o processamento numérico necessário.

A criação dos modelos preditivos foi feita com a biblioteca Scikit-learn, utilizando os algoritmos Random Forest Regressor e MLP Regressor (Rede Neural). A própria biblioteca também foi utilizada para realizar as análises estatísticas de validação, permitindo avaliar o desempenho de cada modelo de forma integrada.

Por fim, a biblioteca Matplotlib foi utilizada para a criação de gráficos. Essa ferramenta permitiu a visualização dos resultados e a comparação direta entre os dados reais de saturação e os valores previstos pelos modelos de inteligência artificial.

4.6 ANÁLISES ESTATÍSTICAS

As métricas consideradas estão descritas a seguir e foram calculadas com base nos dados de teste.

- I. O R^2 , ou coeficiente de determinação, indica o quanto da variação dos dados a variável dependente consegue explicar a partir do modelo. Seu valor varia entre 0 e 1, sendo que quanto mais próximo de 1, melhor o modelo representa os dados. Um valor de R^2 igual a 0 significa que o modelo não explica nenhuma variação, enquanto um valor igual a 1 indica que ele explica toda a variação observada;
- II. O MAE (*Mean Absolute Error*) representa o erro absoluto médio entre os valores reais e os previstos. Ele mostra, em média, o quanto as previsões estão distantes dos valores observados, sem considerar o sentido do erro (positivo ou negativo). É uma métrica simples e fácil de interpretar, pois mantém a unidade da variável de interesse;
- III. O MAPE (*Mean Absolute Percentage Error*) expressa o erro percentual médio absoluto. Ele mostra, em média, a porcentagem de erro em relação aos valores reais, o que facilita a comparação entre diferentes modelos ou contextos. No entanto, o MAPE pode ser distorcido quando os valores reais são muito próximos de zero;
- IV. O RMSE (*Root Mean Squared Error*) é a raiz quadrada da média dos erros ao quadrado. Diferentemente do MAE, o RMSE penaliza erros maiores de forma mais intensa, o que o torna especialmente útil quando grandes desvios são indesejáveis. Assim como o MAE, o RMSE também é expresso na mesma unidade da variável prevista.

5 RESULTADOS E DISCUSSÃO

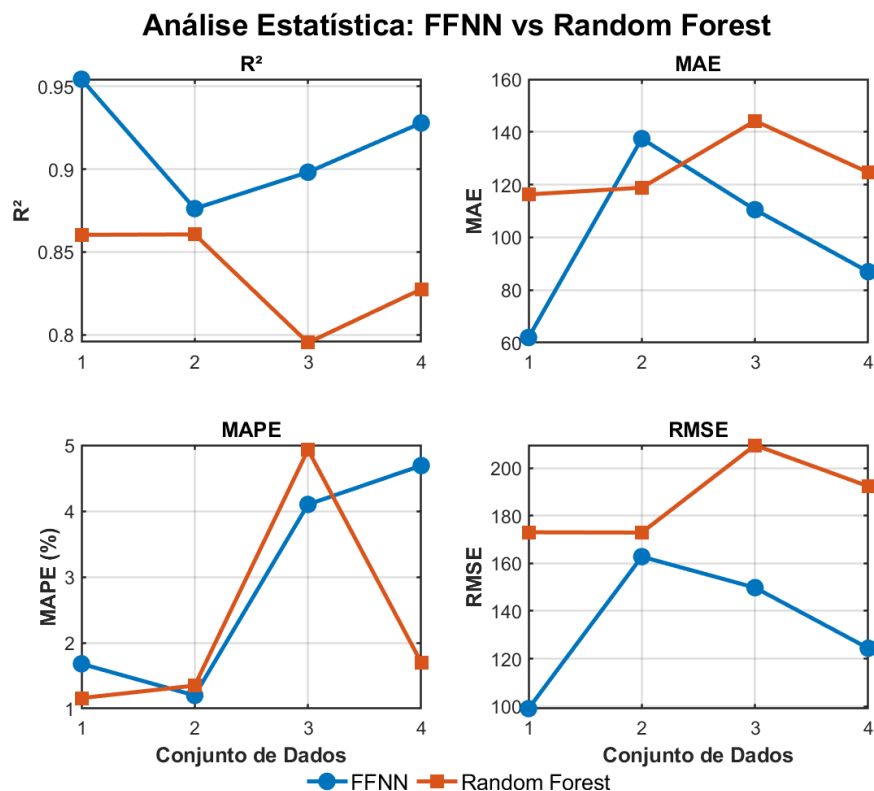
5.1 MODELAGEM

Os testes realizados foram feitos com conjuntos de anos distintos para dados atuais, a fim de analisar a diferença dos resultados encontrados e a relevância de se treinar os modelos com dados mais recentes. As análises foram focadas nos conjuntos 1 e 4 devido ao seu melhor ajuste se comparados aos conjuntos 2 e 3.

A Figura 7 apresenta o desempenho dos modelos testados: MLP e RF, aplicados aos diferentes conjuntos de dados (C1, C2, C3 e C4). Analisando os gráficos da Figura 7, observa-se que o modelo de RF apresentou desempenho visivelmente inferior aos demais. Seu R^2 permaneceu abaixo de 0,90 em quase todos os conjuntos. No C2, houve leve melhora, com R^2 próximo de 0,87, mas os erros MAE e RMSE ainda se mantiveram elevados.

O modelo MLP apresentou o melhor desempenho geral. Os valores de R^2 foram superiores a 0,88 em todos os conjuntos, porém com MAPE variando significativamente com os conjuntos C3 e C4, e, RMSE e MAE tendo um pior resultado com C2. Mesmo sendo mais desafiador em C2 o R^2 ficou próximo de 0,9, demonstrando robustez e boa capacidade de ajuste.

Figura 7: Métricas de desempenho por modelo e conjunto.



Fonte: Próprio autor (2025).

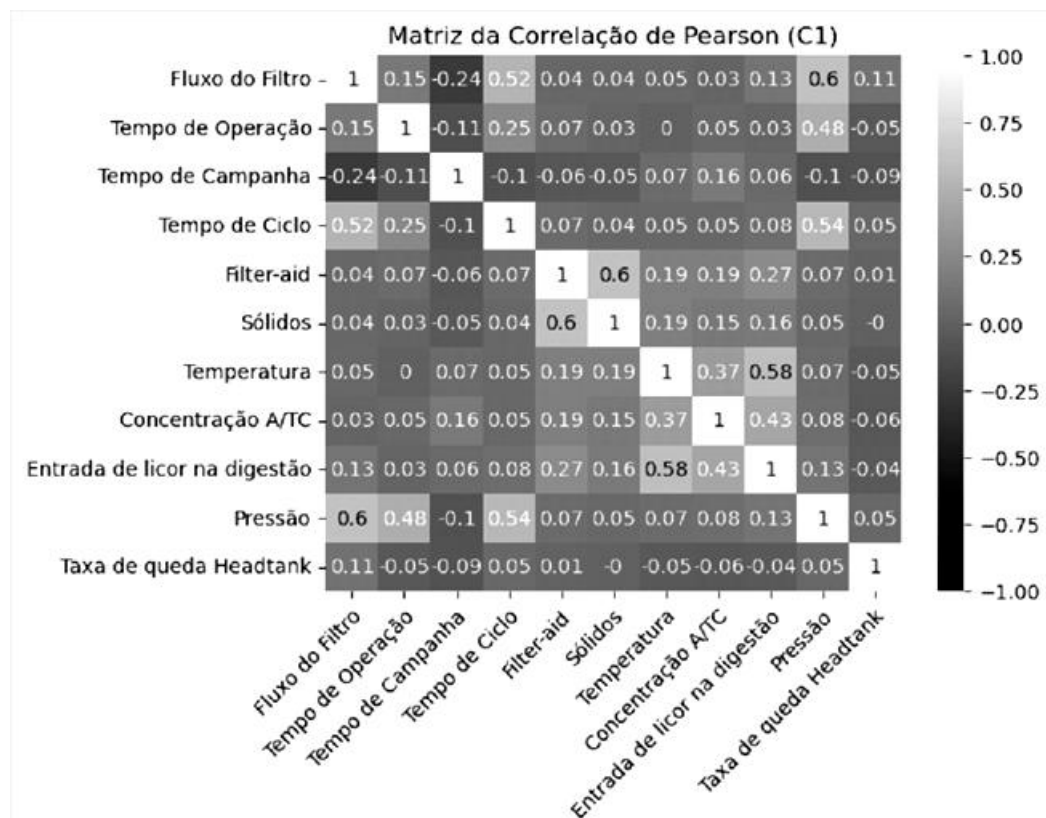
Comparando apenas C1, que obteve o melhor desempenho entre os conjuntos em todos os modelos testados, é possível comparar a dispersão dos valores previstos em relação aos valores reais de cada modelo.

5.2 ANÁLISE DAS CORRELAÇÕES

5.2.1 Correlação de Pearson

A Figura 8 apresenta a matriz de correlação de Pearson referente ao C1, formado pelos dados do ano de 2024. Observa-se uma forte interação entre variáveis operacionais do processo. Dentre as maiores correlações tem-se as variáveis voltadas ao licor como sólidos, temperatura, concentração e entrada de licor na digestão, todas tendo uma ligação maior entre si. Isso se dá devido ao sistema de controle da refinaria atuar diretamente nesses parâmetros, permitindo assim, que se estabeleça relações entre essas variáveis. Além disso, as relações entre pressão com fluxo de filtro, tempo de operação e tempo de ciclo, apesar de não serem variáveis controladas, exibem uma relação linear clara do próprio funcionamento do filtro no decorrer do tempo. De forma geral, a Figura 8 revela um sistema coeso e bem controlado, com respostas operacionais consistentes e interdependência clara entre variáveis-chave.

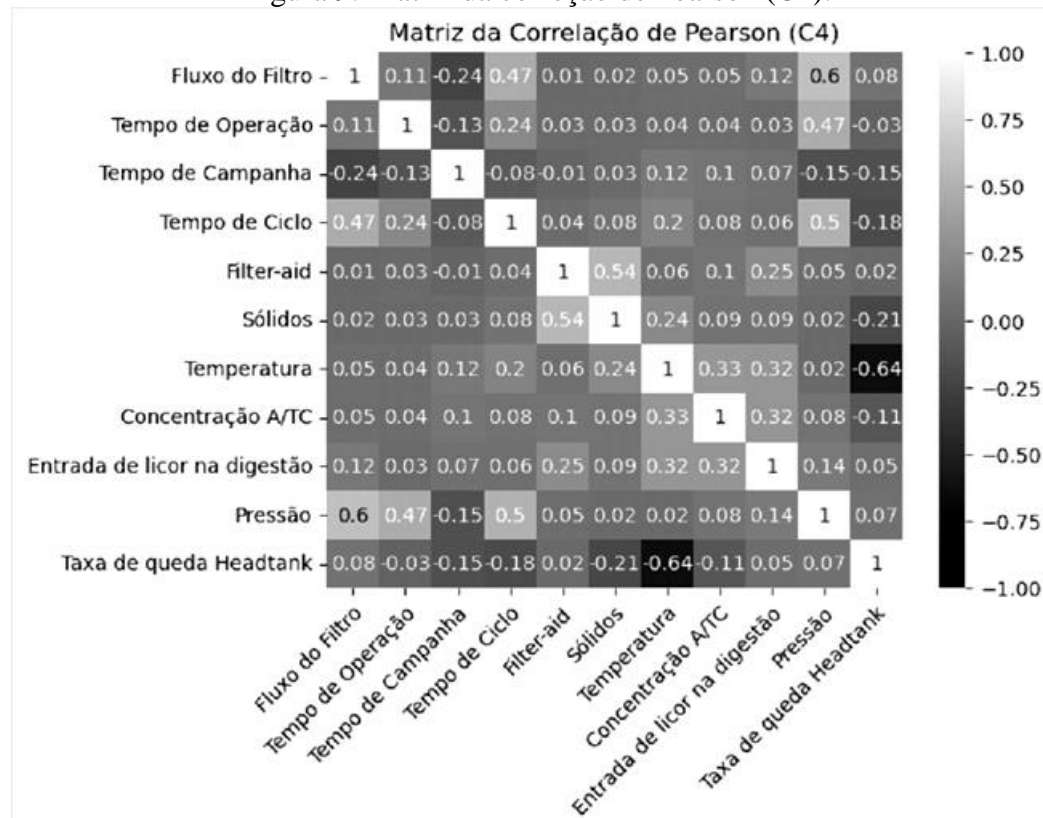
Figura 8: Matriz da correção de Pearson (C1)



Fonte: Próprio autor (2025).

A Figura 9 apresenta a matriz de correlação de Pearson para o C4, formado por dados do período de 2023 a 2024. Embora as mesmas relações estruturais entre variáveis estejam presentes, nota-se que os coeficientes de correlação são, em média, mais baixos do que em C1. As correlações vistas em C4 ainda existem, mas demonstram menor intensidade, o que indica maior variabilidade operacional ao longo do tempo. Essa diferença é esperada em séries históricas mais amplas, onde diferentes campanhas, mudanças de malha, oscilações de carga e ajustes operacionais impactam o alinhamento entre as variáveis. Ainda assim, a Figura 9 confirma que a lógica operacional do processo permanece consistente, mesmo diante de condições variadas, o que reforça a tentativa da operação em manter a estabilidade do processo, mas com muitas flutuações.

Figura 9: Matriz da correção de Pearson (C4).



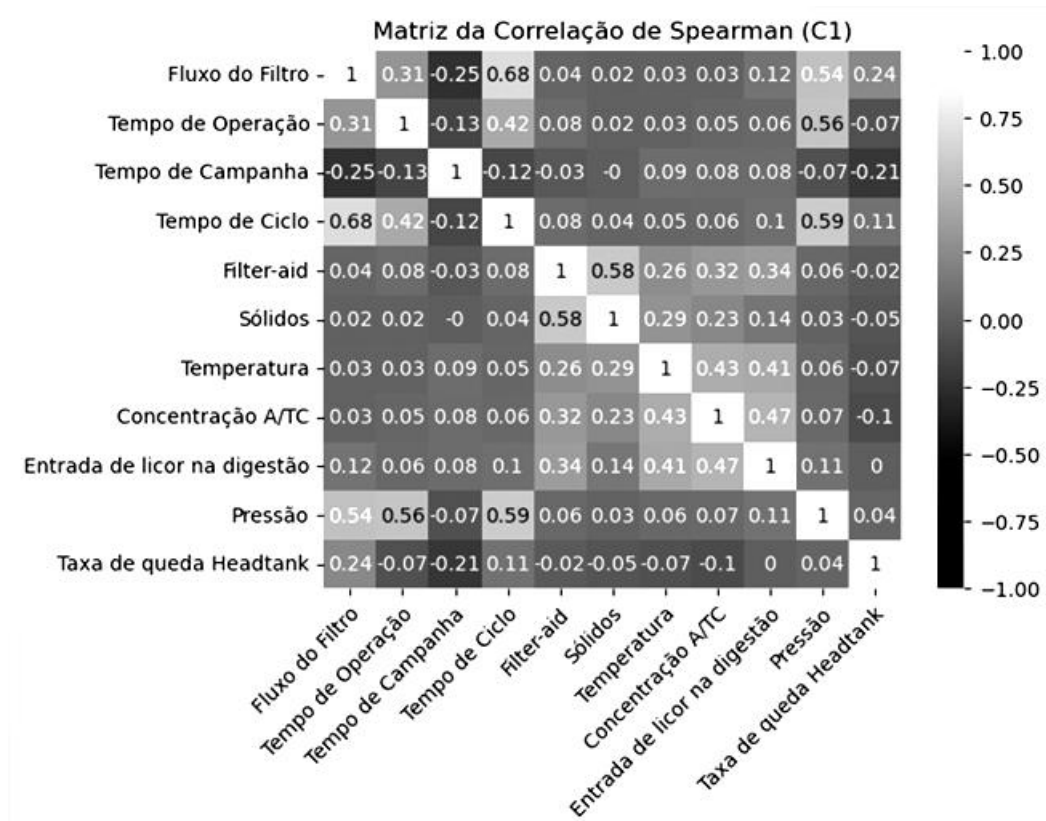
Fonte: Próprio autor (2025).

5.2.2 Correlação de Spearman

A Figura 10 apresenta a matriz de correlação de Spearman para o C1 (dados do ano de 2024). Nela, observam-se relações semelhantes às observadas na correlação de Pearson, porém com algumas alterações. Destaca-se uma relação mais forte das variáveis tempo de ciclo e operação e sua relação com a pressão do filtro, além de um aumento da relação do fluxo do

filtro com o tempo de operação e taxa de queda do *Headtank*. Tais resultados revelam a correlação entre essas variáveis no ano analisado. Algo atípico é que em ambas correlações o tempo de campanha não teve relação com o fluxo do filtro. Esperava-se a queda do fluxo à medida que ao longo do tempo de operação do filtro.

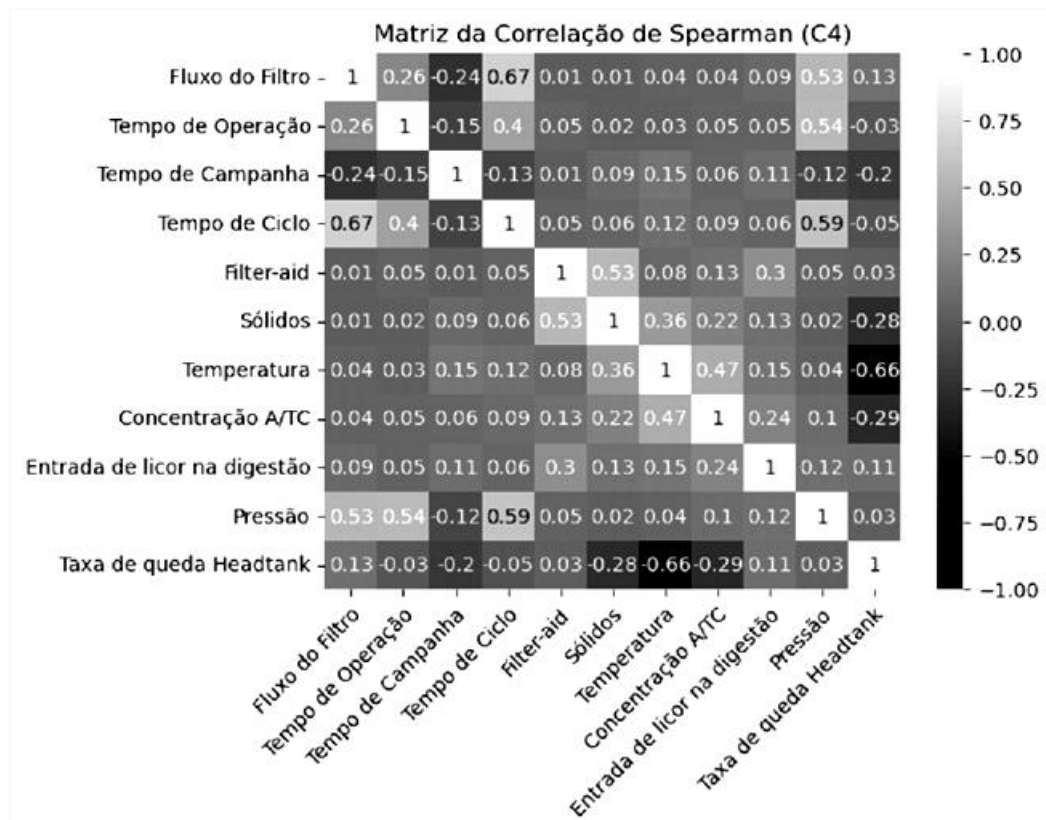
Figura 10: Matriz da correção de Spearman (C1).



Fonte: Próprio autor (2025).

A Figura 11 apresenta a matriz de Spearman para o C4, que compreende os dados do período de 2023 a 2024. Assim como em C1, nota-se a manutenção de boas correlações entre as variáveis analisadas. Um ponto importante nas correlações que utilizaram o conjunto C4, é a queda brusca na variável da temperatura, o que pode significar uma variância na medição do instrumento em um período específico do ano de 2023, que não ocorreu no ano de 2024, podendo tornar o conjunto C4 menos preciso à C1 apesar de abranger um período maior e uma variância de cenários mais abrangente da refinaria.

Figura 11: Matriz da correção de Spearman (C4).

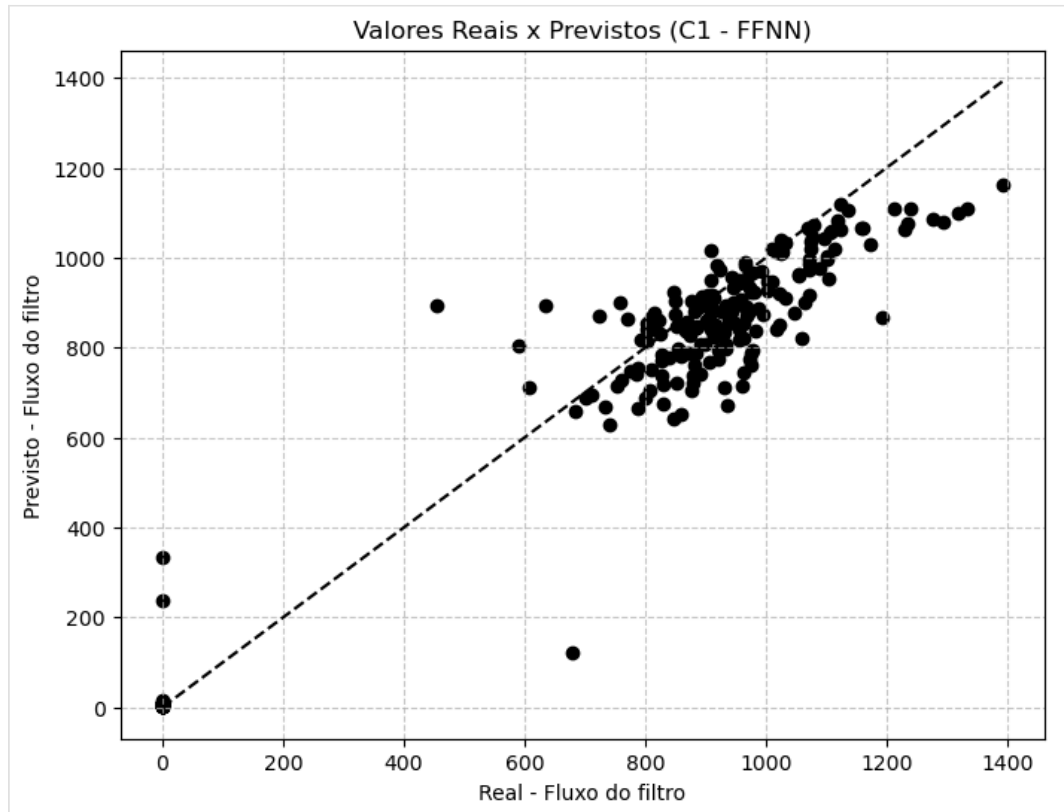


Fonte: Próprio autor (2025).

5.3 COMPARAÇÃO ENTRE VALORES REAIS E PREDITOS

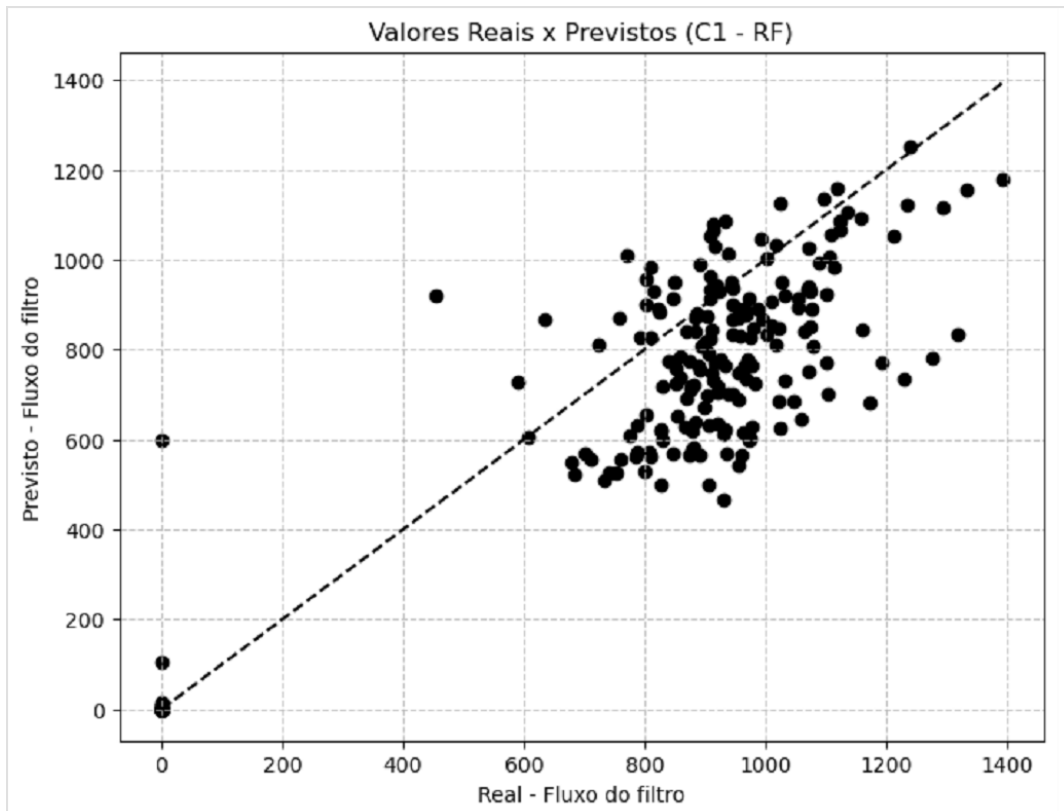
A Figura 12 apresenta o gráfico de dispersão entre os valores reais e os valores previstos pelo modelo MLP aplicado ao C1. Observa-se que os pontos estão concentrados ao longo da linha diagonal, com uma certa tendência para o lado do valor real, indicando que o modelo foi efetivo para previsão dos valores, ainda que com níveis ligeiramente abaixo dos valores reais encontrados, mas em geral, demonstrando uma ótima capacidade preditiva. Parte deste desvio pode ser devido a rotina experimental de realizar purgas manuais nos filtros a fim de forçar um desprendimento da torta quando o filtro está ainda distante do fim de sua operação. A distribuição visual mostra baixa dispersão, com poucos pontos afastados da linha de tendência, o que reforça a precisão do modelo na predição da variável.

Figura 12: Fluxo do filtro: valor previsto versus valor real com MLP no C1.



Fonte: Próprio autor (2025).

Figura 13: Fluxo do filtro: valor previsto versus valor real com RF no C1.



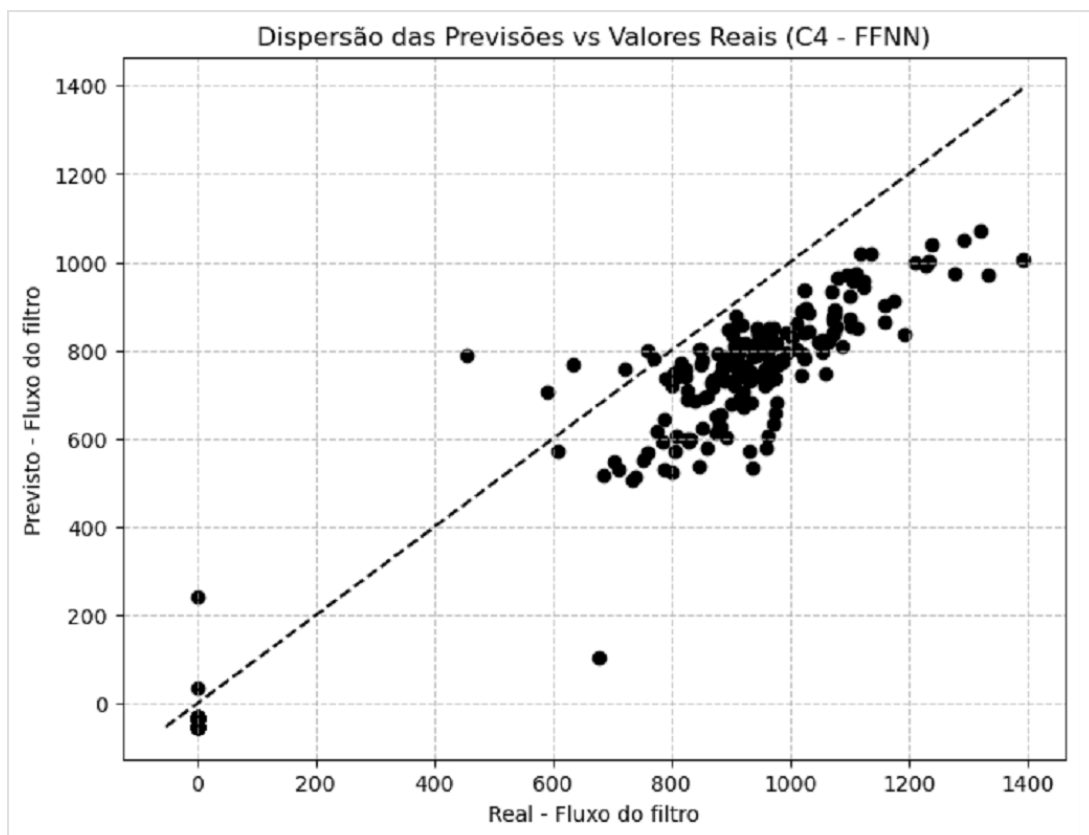
Fonte: Próprio autor (2025).

A Figura 13 apresenta os resultados do modelo RF aplicado ao C1, com o gráfico de dispersão entre os valores reais e os valores previstos. Os pontos apresentam um comportamento semelhante à Figura 12, porém com uma maior dispersão quando comparada ao MLP. A maior dispersão observada em alguns trechos pode estar associada à complexidade do modelo ou à variabilidade natural dos dados, mas o comportamento geral ainda é positivo.

A Figura 14 exibe a dispersão entre os valores reais e os previstos para o modelo MLP no C4. Embora o desempenho não tenha superado o modelo anterior observou-se uma menor dispersão dos dados ao mesmo tempo que o erro tendeu mais para os valores reais quando comparado com C1. Já a Figura 15, exibe a dispersão para o modelo RF no C4. Os resultados foram similares, com um leve aumento na variância dos pontos se comparado ao mesmo modelo no C1.

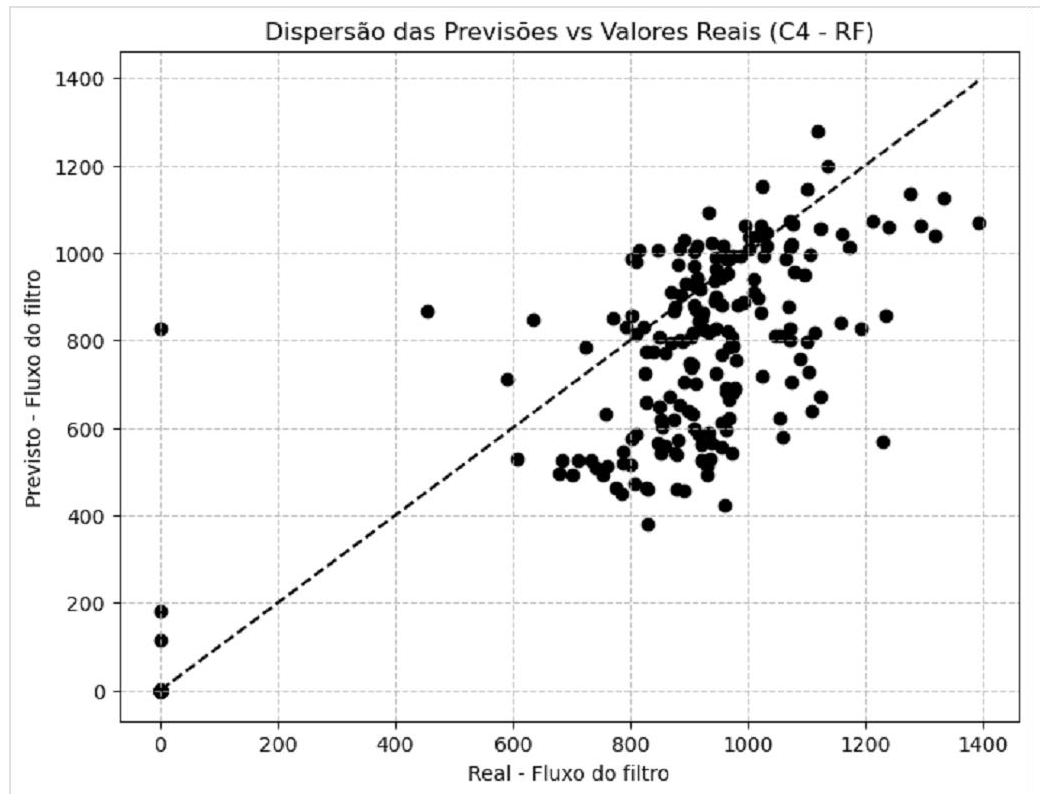
Em todas as dispersões tem-se alguns pontos previstos quando o real estava zerado. Parte disso se dá por alterações bruscas de variáveis intrínsecas ao licor em um período em que o filtro estava inativo, gerando uma falsa predição em alguns pontos no decorrer do tempo.

Figura 14: Fluxo do filtro: valor previsto versus valor real com MLP no C4.



Fonte: Próprio autor (2025).

Figura 15: Fluxo do filtro: valor previsto versus valor real com RF no C4.



Fonte: Próprio autor (2025).

Após todas as análises estatísticas, iniciou-se a criação do modelo para todos os filtros à base do MLP com o conjunto de dados C1.

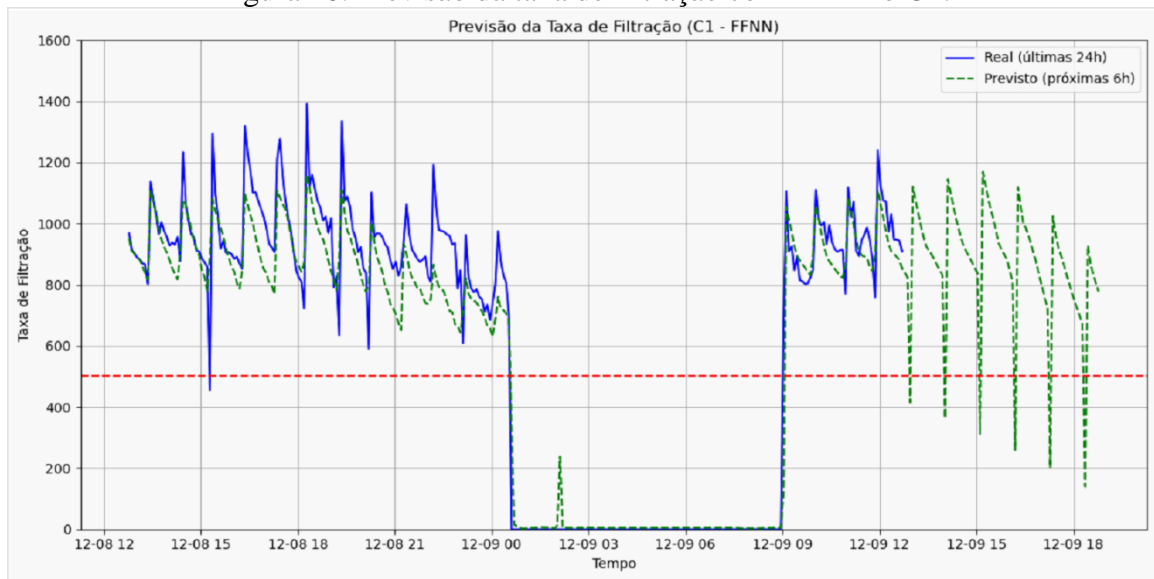
5.4 ANÁLISE E SELEÇÃO DO MODELO PREDITIVO FINAL

Com o melhor modelo e conjunto selecionado, iniciou-se a aplicação com dados em tempo real na planta. Os filtros são divididos em bancadas par e ímpar com numerações de 31 a 40. Os exemplos abaixo focaram apenas em uma bancada para exemplificar o comportamento final do modelo em uma seção inteira do prédio da filtração na refinaria.

O modelo utilizou todas as variáveis listadas anteriormente na tabela 1 ao longo das últimas 24 horas do tempo t , sendo capaz de realizar a predição dos dados, além de aplicar uma média móvel dos mesmos ao longo das próximas 6 horas a partir do tempo t , alcançando assim, uma predição total de 30 horas de funcionamento do filtro (24 horas + 6 horas).

A Figura 16 mostra um exemplo do modelo em funcionamento, sendo a curva azul os valores reais realizado das últimas 24 horas, a curva verde o resultado do modelo e a vermelha o limite de saturação do filtro. No eixo x tem-se o tempo e no eixo y tem-se a nomeado de taxa de filtração o fluxo total do filtro em m^3/h .

Figura 16: Previsão da taxa de filtração com MLP no C1.



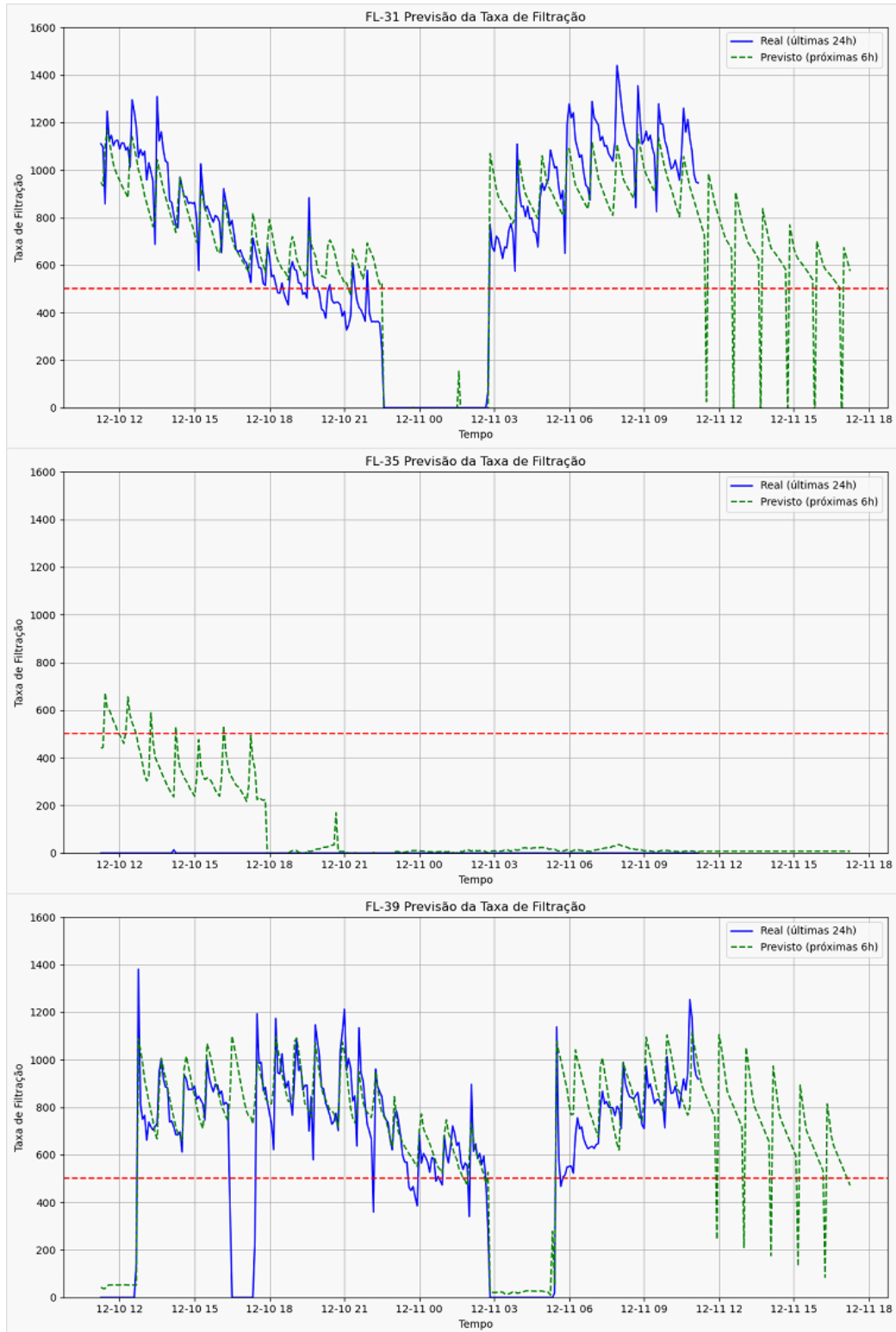
Fonte: Próprio autor (2025).

Adicionalmente, criou-se um alarme para quando o fluxo atingisse a saturação, neste caso 500 m³/h. O alarme funciona como um limite de operação do filtro, indicando quantas horas de operação ele ainda tem pela frente, tornando possível o planejamento de retirada antecipada desse filtro específico de operação para a entrada de outro, assim como, permitindo a preparação antecipada do próximo filtro que substituirá o saturado.

O sistema atua como um constante tempo ótimo de operação, possibilitando a organização dos tempos de operação e lavagem de cada filtro, sem forçar ou desgastar desnecessariamente os equipamentos envolvidos na operação, além de diminuir perdas de produção por paradas indesejadas.

A Figura 17 exemplifica o modelo completo prevendo uma parte da bancada ímpar de filtros. É possível observar um bom ajuste do modelo aos dados, assim como uma previsão do filtro 35, que por algum problema técnico na área, não está operando de forma automática. Isso se dá porque, com exceção do fluxo, todas as variáveis estão sendo medidas no campo através dos outros instrumentos, possibilitando a previsão até quando o filtro está *offline*.

Figura 17: Previsão da taxa de filtração para filtros 31, 35 e 39.

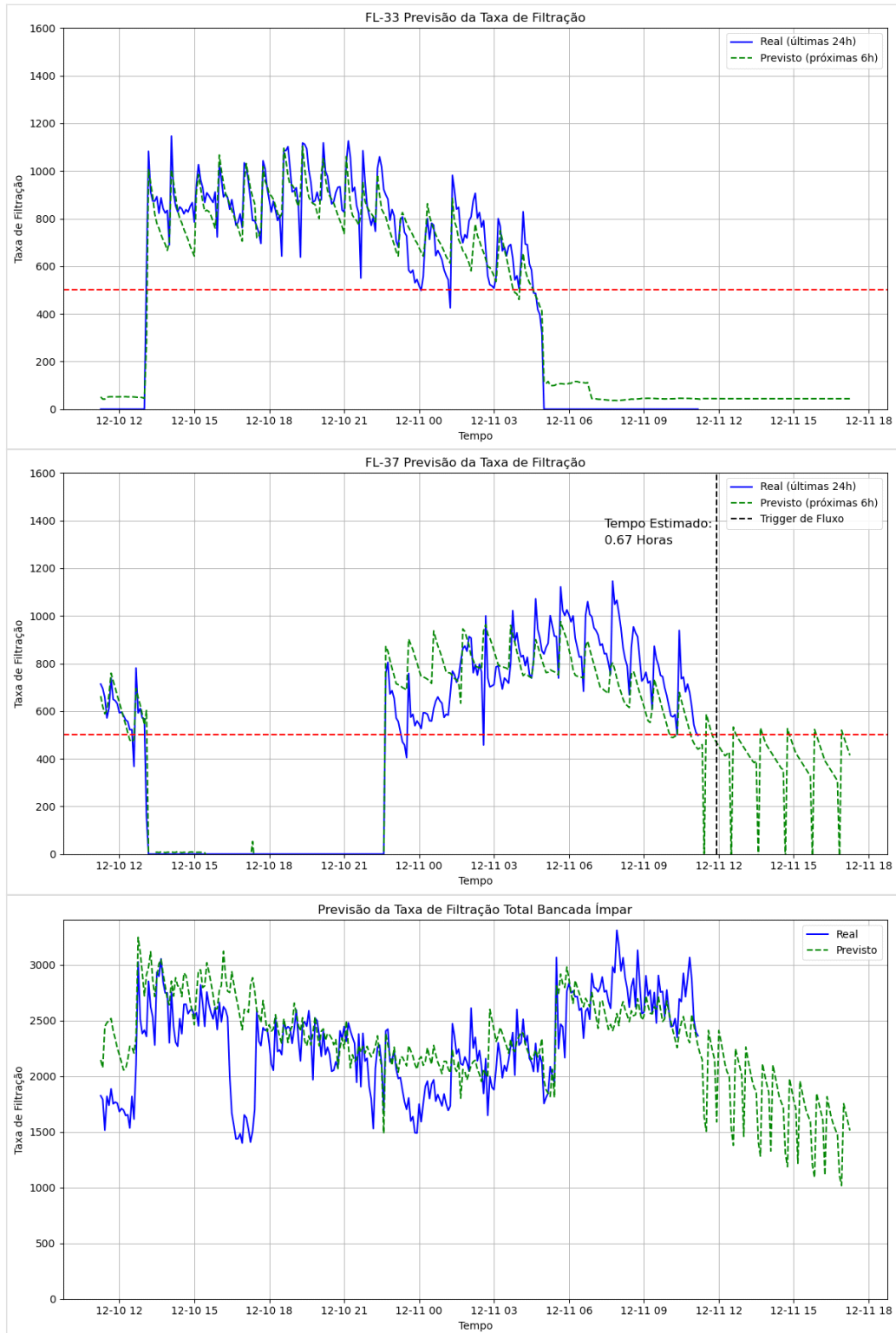


Fonte: Próprio autor (2025).

Na figura 18, que contém o restante da bancada em conjunto com o fluxo total da bancada, é possível observar o mesmo tipo de resultado encontrado na figura 17 e também o alarme mencionado no filtro 37, indicando uma saturação em menos de 1 hora.

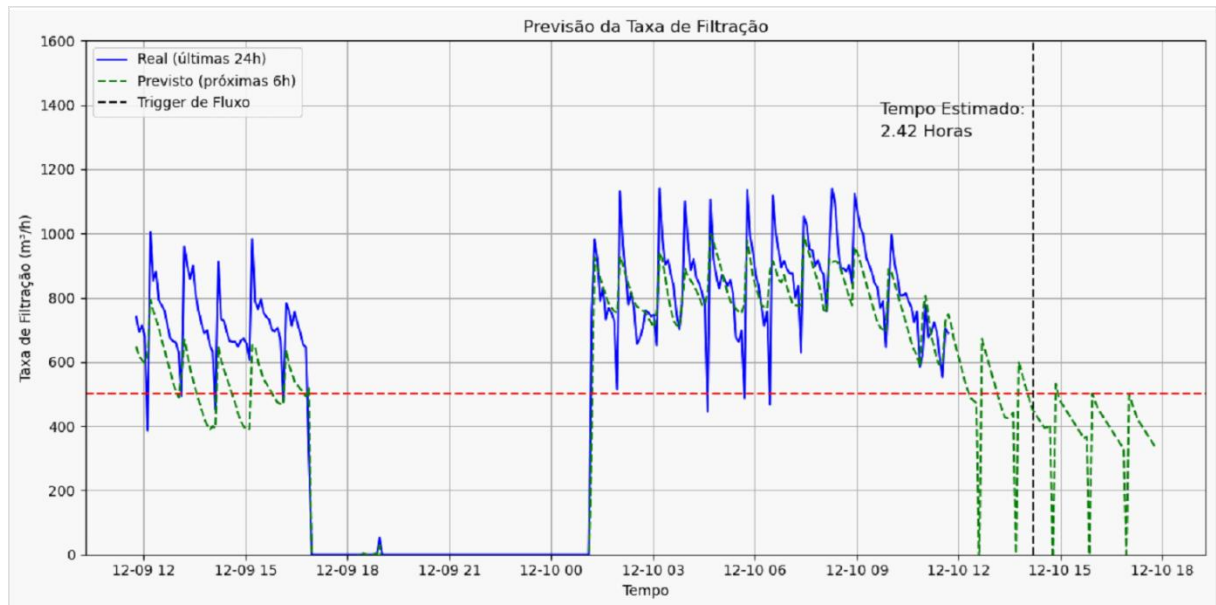
A Figura 19 por sua vez, exemplifica o modelo em funcionamento com mais detalhes, sendo possível observar o alerta de saturação cerca de duas horas e meia antes que ocorra, o que permite tempo suficiente para que haja a programação da lavagem de um filtro para substituição.

Figura 18: Previsão da taxa de filtração para os filtros 33, 37 e o total da bancada.



Fonte: Próprio autor (2025).

Figura 19: Previsão da taxa de filtração para um de filtro com alarme de saturação.



Fonte: Próprio autor (2025).

A tabela 3 mostra o desempenho dos modelos em diferentes configurações para C1 e seus respectivos R^2 . Como resultado final, tem-se uma rede neural MLP com duas camadas ocultas de 20 e 10 neurônios respectivamente, utilizando-se da função de ativação ReLu e otimizador Adam como o melhor modelo preditivo encontrado para os filtros.

Tabela 3: Resultados obtidos em C1.

Rank	Modelo	Configuração Principal	R2
1	MLP (Rede Neural)	2 camadas (20, 10), ReLU, Adam	0.96
2	MLP (Rede Neural)	2 camadas (10, 5), ReLu, Adam	0.89
3	Random Forest	20 árvores, Profundidade máx: 15	0.86
4	Random Forest	15 árvores, Profundidade máx: 10	0.80

Fonte: Próprio autor (2025).

6 CONCLUSÃO

- A análise comparativa de modelos revelou que as abordagens de aprendizado de máquina foram capazes de prever o comportamento dos filtros em diversos cenários, o que é relevante, considerando-se a complexidade e a variabilidade dos dados do processo, em que as relações entre as variáveis podem não ser puramente lineares.
- A rede neural se mostrou mais eficaz na predição, principalmente pelo fato possibilitar um ajuste mais fino à não linearidade específica dos dados do processo de filtração.
- A aplicação da técnica de normalização Min-Max, aliada à função de ativação ReLu, se mostraram, não só adequadas aos dados, como facilitaram a convergência e manipulação do modelo como um todo.
- As análises das variáveis evidenciam que algumas tendências são sazonais ou pontuais, expondo o impacto de ajustes operacionais específicos ao longo dos anos e a necessidade de manter modelos constantemente treinados com dados mais recentes de operação da planta.

REFERÊNCIAS

- BACH, Manfred; SCHOENBRUNN, Fred. Liquid–Solid Separation: De–Sanding—Flocculation, Sedimentation and Liquor Filtration. *In: [S.l.: S.n.]*. p. 241–267.
- BOTT, Reinhard; LANGELOH, Thomas; HAHN, Jürgen. Advanced Filtration Methods for Pregnant Liquor Purification. *In: Essential Readings in Light Metals*. Cham: Springer International Publishing, 2016. p. 444–448.
- BREIMAN, L. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, v. 16, n. 3, p. 199–231, ago. 2001.
- BREIMAN, Leo. *Random Forests*. *Machine Learning*, v. 45, n. 1, p. 5–32, out. 2001.
- BUTEAN, Alex *et al.* A Review of Artificial Intelligence Applications for Biorefineries and Bioprocessing: From Data-Driven Processes to Optimization Strategies and Real-Time Control. *Processes*, v. 13, n. 8, p. 2544, 12 ago. 2025.
- CHENG, Lu-wei *et al.* Scale Formation During the Bayer Process and a Potential Prevention Strategy. *Journal of Sustainable Metallurgy*, v. 7, n. 3, p. 1293–1303, 17 set. 2021.
- COOLING, D. J. Developments in the Disposal of Residue from the Alumina Refining Industry. *In: Essential Readings in Light Metals*. Cham: Springer International Publishing, 2016. p. 927–932.
- DOBBELAERE, Maarten R. *et al.* Machine Learning in Chemical Engineering: Strengths, Weaknesses, Opportunities, and Threats. *Engineering*, v. 7, n. 9, p. 1201–1211, set. 2021.
- DUTTA, Souradeep *et al.* Output Range Analysis for Deep Feedforward Neural Networks. *In: [S.l.: S.n.]*. p. 121–138.
- GONTIJO, Glayson Stopa *et al.* Improving the Bayer Process productivity – An industrial case study. *Minerals Engineering*, v. 22, n. 13, p. 1130–1136, out. 2009.
- HABASHI, Fathi. A Hundred Years of the Bayer Process for Alumina Production. *In: Essential Readings in Light Metals*. [S.l.]: Wiley, 2013. p. 83–93.
- HIND, Andrew R.; BHARGAVA, Suresh K.; GROCCOTT, Stephen C. The surface chemistry of Bayer process solids: a review. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, v. 146, n. 1–3, p. 359–374, jan. 1999.
- KAMAL ABDELBASSET, Walid *et al.* Development of multiple machine-learning computational techniques for optimization of heterogenous catalytic biodiesel production from waste vegetable oil. *Arabian Journal of Chemistry*, v. 15, n. 6, p. 103843, jun. 2022.
- KOMLÓSSY, George; VAN DEURSEN, Caio; RAAHAUGE, Benny E. Bauxite: Geology, Mineralogy, Resources, Reserves and Beneficiation. *In: [S.l.: S.n.]*. p. 19–132.
- LAMACCHIA, Robert; COSTA, Raphael. Bayer Process—Water and Energy Balance. *In: [S.l.: S.n.]*. p. 427–504.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. *Nature*, v. 521, n. 7553, p. 436–444, 28 maio 2015.

LI, Ying *et al.* Multi-scale revolution of artificial intelligence in chemical industry. *Frontiers of Chemical Science and Engineering*, v. 19, n. 7, p. 57, 23 jul. 2025.

MARKETGROUP. Filter Bags. Disponível em: <<https://www.markert-group.com/en/textile-filtration/filterbags>>. Acesso em: 10 dez. 2025.

MELO, Caio *et al.* Enhancing Solid-Liquid Separation in Alumina Refineries – A Comparative Study of Gravity Decanters and *Diastar* Filters. *In: Dubai: ICSOBA*, 6 out. 2023.

PATIL, Sharad V.; THORAT, Bhaskar N. Mechanical dewatering of red mud. *Separation and Purification Technology*, v. 294, p. 121157, ago. 2022.

RYNKIEWICZ, J. General bound of overfitting for MLP regression models. *Neurocomputing*, v. 90, p. 106–110, ago. 2012.

SAAIDI, Afaf; BICHRI, Ahmed; ABDERAFI, Souad. Efficient machine learning model to predict dynamic viscosity in phosphoric acid production. *Results in Engineering*, v. 18, p. 101024, jun. 2023.

SARKER, Iqbal H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, v. 2, n. 3, p. 160, 22 maio 2021.

SCHONLAU, Matthias; ZOU, Rosie Yuyan. The *Random Forest* algorithm for statistical learning. *The Stata Journal: Promoting communications on statistics and Stata*, v. 20, n. 1, p. 3–29, 24 mar. 2020.

SHAW, Ruth; DUNCAN, Ajamu; CROSDALE, Marlon. Improvements in Smelter Grade Alumina Quality at Clarendon Alumina Works. *In: Light Metals 2011. [S.l.]: Wiley*, 2011. p. 75–77.

SINGARIMBUN, Roy Nuary; NABABAN, Erna Budhiarti; SITOMPUL, Opim Salim. Adaptive Moment Estimation To Minimize Square Error In Backpropagation Algorithm. *In: IEEE*, nov. 2019.

SINGH, Dalwinder; SINGH, Birmohan. Investigating the impact of data normalization on classification performance. *Applied Soft Computing*, v. 97, p. 105524, dez. 2020.

SZANDAŁA, Tomasz. Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. *In: [S.l.: S.n.]*. p. 203–224.

THIRUMALAI, Chandrasegar; CHANDHINI, Swapna Anupriya; VAISHNAVI, M. Analysing the concrete compressive strength using Pearson and Spearman. *In: IEEE*, abr. 2017.