

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ

**CELSO HAYASHI SINZATO JUNIOR
PEDRO MENARIM MOERS**

**OTIMIZAÇÃO DE CONTROLADORES PID E FUZZY-PID EM SISTEMAS
DE REGULAÇÃO TÉRMICA COM APRENDIZAGEM DE MÁQUINA**

PONTA GROSSA

2025

**CELSO HAYASHI SINZATO JUNIOR
PEDRO MENARIM MOERS**

**OTIMIZAÇÃO DE CONTROLADORES PID E FUZZY-PID EM SISTEMAS
DE REGULAÇÃO TÉRMICA COM APRENDIZAGEM DE MÁQUINA**

**Optimization of PID and Fuzzy-PID controllers in
thermal regulation systems with machine learning**

Trabalho de conclusão de curso de graduação
apresentado como requisito para obtenção do título de
Bacharel em Engenharia Elétrica da Universidade
Tecnológica Federal do Paraná (UTFPR).

Orientadora:
Prof^a Dra. Marcella Scoczynski Ribeiro Martins

PONTA GROSSA

2025



Esta licença permite download e compartilhamento do trabalho desde que sejam atribuídos créditos ao(s) autor(es), sem a possibilidade de alterá-lo ou utilizá-lo para fins comerciais. Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**CELSO HAYASHI SINZATO JUNIOR
PEDRO MENARIM MOERS**

**OTIMIZAÇÃO DE CONTROLADORES PID E FUZZY-PID EM SISTEMAS
DE REGULAÇÃO TÉRMICA COM APRENDIZAGEM DE MÁQUINA**

Trabalho de conclusão de curso de graduação
apresentado como requisito para obtenção do título de
Bacharel em Engenharia Elétrica da Universidade
Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 25/novembro/2025

Marcella Scoczynski Ribeiro Martins
Doutorado
Universidade Tecnológica Federal do Paraná

Edison Luiz Salgado Silva
Mestrado
Universidade Tecnológica Federal do Paraná

Cristhiane Goncalves
Doutorado
Universidade Tecnológica Federal do Paraná

**PONTA GROSSA
2025**

RESUMO

Esta pesquisa insere-se no contexto da Indústria 4.0, em que a precisão e a eficiência de sistemas de controle térmico tornam-se cada vez mais essenciais. O trabalho tem como objetivo aplicar e avaliar diferentes modelos de Aprendizagem de Máquina para a otimização da ação de controle em sistemas baseados em controladores Proporcional-Integral-Derivativo (PID) e Fuzzy-PID. A metodologia seguiu um fluxo de Aprendizagem Supervisionada voltado à regressão, utilizando uma base de dados pública proveniente de um sistema de manufatura inteligente. Após a etapa de preparação dos dados — incluindo limpeza, padronização e seleção das variáveis — foram treinados e avaliados nove modelos de regressão, entre eles Random Forest, Multi-Layer Perceptron e diferentes variantes de Regressão Linear (Linear, Ridge e Lasso), SVR, Gradient Boosting, XGBoost e LightGBM. O alvo previsto pelos modelos corresponde à ação de controle fornecida pelo controlador Fuzzy-PID. Os resultados demonstram que a Aprendizagem de Máquina pode capturar de forma eficaz o comportamento do controlador adaptativo, mesmo em sistemas térmicos não lineares e sujeitos a histerese e variações ambientais. Entre os modelos testados, a Regressão Lasso apresentou o melhor desempenho global, alcançando $R^2 = 0,9407$ e os menores erros médios, indicando que a relação entre variáveis de processo e ação de controle segue uma estrutura suficientemente regular para ser aprendida por um modelo linear regularizado. Esses achados reforçam o potencial da Aprendizagem de Máquina como ferramenta de apoio à sintonia de controladores, oferecendo maior previsibilidade, eficiência computacional e redução do esforço de calibração. O estudo evidencia que técnicas modernas de regressão podem complementar — e, em alguns cenários, superar — abordagens tradicionais, contribuindo diretamente para o avanço de soluções inteligentes em ambientes industriais alinhados aos princípios da Indústria 4.0.

Palavras-chave: regressão Lasso; controlador PID; lógica Fuzzy; aprendizagem de Máquina; regulação térmica.

ABSTRACT

This research is situated within the scope of Industry 4.0, where accuracy and efficiency in thermal control systems have become increasingly essential. The main objective of this work is to apply and evaluate different Machine Learning models to optimize the control action in systems based on Proportional-Integral-Derivative (PID) and Fuzzy-PID controllers. The methodology followed a supervised learning workflow focused on regression, using a public dataset extracted from an intelligent manufacturing system. After data preparation—including cleaning, standardization, and feature selection—nine regression models were trained and evaluated, such as Random Forest, Multi-Layer Perceptron, and different variants of Linear Regression (Linear, Ridge, and Lasso), SVR, Gradient Boosting, XGBoost and LightGBM. The prediction target corresponded to the control action generated by the Fuzzy-PID controller. The results show that Machine Learning can effectively capture the behavior of the adaptive controller, even in nonlinear thermal systems subject to hysteresis and environmental variations. Among the models tested, Lasso Regression achieved the best overall performance, reaching $R^2 = 0.9407$ and the lowest error metrics, indicating that the relationship between process variables and control action follows a sufficiently regular structure to be learned by a regularized linear model. These findings reinforce the potential of Machine Learning as a support tool for controller tuning, offering greater predictability, computational efficiency, and reduced calibration effort. The study highlights that modern regression techniques can complement—and in some scenarios, outperform—traditional approaches, directly contributing to the advancement of intelligent solutions in industrial environments aligned with Industry 4.0 principles.

Keywords: Lasso regression; PID controller; Fuzzy logic; machine learning; thermal regulation.

SUMÁRIO

1	INTRODUÇÃO	7
1.1	Objetivo Geral	9
1.2	Objetivos Específicos	9
2	REVISÃO DA LITERATURA	10
2.1	PID Clássico	10
2.2	Evolução Histórica e Tecnológica até a Indústria 4.0	11
2.3	Aprendizagem de Máquina	13
2.3.1	Regressão	13
2.3.1.1	Random forest	14
2.3.1.2	Multi-Layer Perception (MLP)	14
2.3.1.3	Regressão Linear	14
2.3.1.4	Regressão de Ridge	15
2.3.1.5	Regressão de Lasso	15
2.3.1.6	SVR	15
2.3.1.7	Gradient Boosting Regressor	16
2.3.1.8	XGBoost	16
2.3.1.9	LightGBM	16
2.3.2	Lógica Fuzzy	17
2.3.3	Metodos avaliativos	17
2.3.3.1	Coefficiente de determinação (R^2)	17
2.3.3.2	Erro Quadrático Médio (MSE)	18
2.3.3.3	Raiz do Erro Quadrático Médio (RMSE)	19
2.3.3.4	Erro Absoluto Médio (MAE)	19
2.3.3.5	Gráfico de Resíduos:	20
2.4	Trabalhos Correlatos	20
3	MATERIAL E MÉTODOS	24
3.1	Base de dados	24
3.1.1	Parâmetros de entrada	25
3.1.2	Parâmetro de saída	26
3.2	Método aplicado	26
3.2.1	Preparação e carregamento de dados	27
3.2.2	Pré-processamento e engenharia de <i>features</i>	27
3.2.3	Treinamento e avaliação inicial dos modelos	28
3.2.4	Modelagem e otimização	28
3.2.5	Avaliação de desempenho	30
4	RESULTADOS E DISCUSSÃO	31
4.1	Coefficiente de Determinação (R^2)	31
4.2	Erro Quadrático Médio (MSE)	32
4.3	Raiz do Erro Quadrático Médio (RMSE)	33
4.4	Erro Absoluto Médio (MAE)	33
4.5	Gráficos de resíduos	34
4.6	Gráficos de saída real versus saída predita	38
4.7	Síntese dos resultados e implicações práticas	41
5	CONCLUSÕES	43

REFERÊNCIAS 45

1 INTRODUÇÃO

A presente pesquisa situa-se no contexto da Indústria 4.0, marco tecnológico caracterizado pela convergência entre sistemas físicos e digitais, automação inteligente e análise de dados em larga escala. Concebido no início do século XXI, esse paradigma transformou os processos produtivos ao integrar sensores, sistemas ciberfísicos, computação em nuvem, Internet das Coisas (IoT) e algoritmos avançados de Inteligência Artificial. Essa sinergia possibilita níveis inéditos de produtividade, flexibilidade, rastreabilidade e autonomia operacional, favorecendo tomadas de decisão rápidas e embasadas em dados (Antunes, 2020).

Nesse cenário, o controle de processos emerge como um dos pilares fundamentais da automação moderna. Em especial, o controle de temperatura exerce papel crítico em diversos segmentos industriais — como manufatura, química, farmacêutica, alimentos, eletrônica e energia — influenciando diretamente eficiência energética, estabilidade do processo, qualidade do produto final e segurança operacional. A precisão nesse controle é essencial para evitar oscilações indesejadas, reduzir desperdícios e garantir conformidade com padrões técnicos rigorosos.

Historicamente, os controladores Proporcional–Integral–Derivativo (PID) são considerados como a estratégia mais difundida na indústria. Sua popularidade deve-se à simplicidade estrutural, baixo custo computacional, interpretabilidade e capacidade de produzir respostas adequadas em uma ampla variedade de sistemas dinâmicos (Ogata, 2010). Entretanto, a eficácia dos controladores PID depende fortemente da sintonia adequada dos ganhos K_p , K_i e K_d . Em processos lineares ou pouco sensíveis a perturbações, métodos clássicos de sintonia (como Ziegler–Nichols ou Lugar das Raízes) atendem de maneira satisfatória. Contudo, à medida que os sistemas tornam-se mais complexos — apresentando não linearidades, histerese, perturbações externas, variações térmicas e incertezas dinâmicas — os controladores PID tradicionais enfrentam limitações significativas.

Nesse contexto, estratégias híbridas, como os controladores Fuzzy-PID, têm sido amplamente estudadas. Ao incorporar regras linguísticas e raciocínio aproximado, esses controladores conferem maior adaptabilidade e robustez ao sistema, permitindo ajustes dinâmicos dos ganhos PID conforme as condições operacionais variam. No entanto, o processo de determinação dessas regras fuzzy ou da função de mapeamento

ideal ainda depende de conhecimento especializado e pode demandar longos ciclos de testes experimentais.

Diante desse cenário, este trabalho investiga a aplicação da Aprendizagem de Máquina (Machine Learning) como ferramenta de apoio à otimização de controladores térmicos. A Aprendizagem de Máquina constitui um campo da Inteligência Artificial dedicado ao desenvolvimento de métodos capazes de aprender padrões a partir de dados, realizar previsões e adaptar-se automaticamente ao ambiente (Monard; Baranauskas, 2003). Em particular, técnicas de regressão (Izbicki; Santos, 2020) são amplamente utilizadas para modelar relações contínuas entre variáveis, sendo adequadas para prever ações de controle em sistemas dinâmicos.

Nesta pesquisa, são aplicados e comparados modelos de regressão supervisionada para estimar diretamente a saída ideal do controlador Fuzzy-PID. Utiliza-se uma base de dados real extraída de um sistema de manufatura inteligente, composta por variáveis térmicas, parâmetros de controle e indicadores de desempenho a cada ciclo de operação. Ao todo, nove modelos foram treinados e avaliados: Regressão Linear, Ridge, Lasso, Random Forest, Multi-Layer Perceptron (MLP), Support Vector Regression (SVR), Gradient Boosting, XGBoost e LightGBM, abrangendo métodos clássicos, regularizados, ensemble e redes neurais.

A avaliação de desempenho baseou-se em métricas amplamente aceitas na literatura de modelagem estatística e regressão: Coeficiente de Determinação (R^2), Erro Quadrático Médio (MSE), Raiz do Erro Quadrático Médio (RMSE), Erro Absoluto Médio (MAE) e análise gráfica de resíduos, assegurando uma validação robusta dos modelos (Montgomery; Peck; Vining, 2021). O objetivo central consiste em identificar qual abordagem oferece a melhor capacidade preditiva para a saída do controle Fuzzy-PID, contribuindo para o desenvolvimento de sistemas de regulação térmica mais eficientes, autônomos e alinhados ao ambiente da Indústria 4.0.

Além disso, o presente estudo busca responder às seguintes questões fundamentais:

- É possível aproximar a lógica de decisão de um controlador Fuzzy-PID através de modelos supervisionados?
- Entre modelos lineares, de ensemble e de redes neurais, qual apresenta maior capacidade preditiva nesse contexto?
- Como o aprendizado baseado em dados pode reduzir o esforço de sintonia

tradicional de controladores térmicos?

- Em que medida a Aprendizagem de Máquina pode contribuir para processos autônomos dentro da Indústria 4.0?

Ao integrar técnicas modernas de Aprendizagem de Máquina ao domínio clássico do controle automático, este trabalho pretende evidenciar caminhos promissores para a automação inteligente, contribuindo para sistemas industriais mais adaptativos, robustos e eficientes.

1.1 Objetivo Geral

Investigar, aplicar e avaliar técnicas de Aprendizagem de Máquina para prever a ação de controle ideal em sistemas térmicos baseados em controladores Fuzzy-PID, de modo a aprimorar o desempenho de processos industriais no contexto da Indústria 4.0.

1.2 Objetivos Específicos

- Preparar e tratar a base de dados proveniente de um sistema de manufatura inteligente, realizando limpeza, padronização, seleção de variáveis e estruturação dos conjuntos de treino e teste.
- Treinar e comparar diferentes modelos de regressão supervisionada — incluindo abordagens lineares, regularizadas, ensemble e redes neurais — para estimar a saída de controle gerada pelo controlador Fuzzy-PID.
- Avaliar o desempenho dos modelos utilizando métricas estatísticas consagradas, tais como R^2 , MSE, RMSE, MAE e análise de resíduos, assegurando uma validação robusta e interpretável.
- Identificar qual modelo apresenta a melhor capacidade preditiva, examinando sua eficiência em capturar padrões complexos presentes no comportamento do controlador Fuzzy-PID.

2 REVISÃO DA LITERATURA

Este capítulo introduz conceitos importantes para o desenvolvimento do trabalho. Serão apresentadas técnicas de PID, abrangendo formas clássicas e também a utilização de novas maneiras de implementações, como o PID Fuzzy. Por fim abordaremos sua utilização na indústria 4.0, bem como técnicas de aprendizagem de máquina.

2.1 PID Clássico

Os controladores Proporcional, Integral e Derivativo (PID) são a base da automação industrial, sendo amplamente utilizados em mais de 90% dos sistemas de controle do mundo (Ogata, 2010). A sua popularidade se deve à sua estrutura simples, robustez e eficácia para uma vasta gama de aplicações. A equação básica de um controlador PID é dada pela equação abaixo:

$$u(t) = K_p e(t) + K_i \int e(t) dt + K_d \frac{d}{dt} e(t) \quad (1)$$

Onde:

- $u(t)$ – é o sinal de controle (saída do controlador);
- $e(t)$ - é o erro (diferença entre o ponto de ajuste e a variável de processo);
- K_p – é o ganho proporcional, que ajusta a resposta atual do erro;
- K_i – é o ganho integral, que elimina o erro em regime permanente;
- K_d – é o ganho derivativo, que prevê o erro futuro com base na taxa de variação;

A sintonia dos parâmetros (K_p , K_i , K_d) é crucial para garantir o desempenho desejado do sistema, como tempo de acomodação rápido, baixo overshoot e erro nulo em regime permanente. Existem diversos métodos de sintonia clássica, sendo os mais notáveis:

- *Métodos Empíricos de Ziegler-Nichols*: Desenvolvidos por John G. Ziegler e Nathaniel B. Nichols, esses métodos são baseados em testes práticos. Eles envolvem a determinação do ganho crítico e do período de oscilação do sistema em malha fechada, a partir dos quais os parâmetros do PID são calculados.

- *Métodos por Análise de Resposta em Frequência:* Incluem as técnicas de Nyquist e Bode, que analisam o comportamento do sistema em diferentes frequências para garantir a estabilidade e a margem de fase.
- *Métodos de Cohen-Coon:* Úteis para sistemas com atraso de tempo, esses métodos utilizam a resposta ao degrau do processo para determinar os parâmetros do controlador.

2.2 Evolução Histórica e Tecnológica até a Indústria 4.0

De acordo com (Sacomano et al., 2018), os primeiros fundamentos do que viria a se tornar a internet surgiram durante a Guerra Fria, por volta de 1957. A iniciativa, desenvolvida pelos Estados Unidos, tinha como objetivo principal descentralizar informações e aumentar a eficiência no envio e recebimento de dados. Em 1969, consolidou-se a primeira rede operacional, baseada na transmissão de dados segmentados em pequenos pacotes que percorriam rotas distintas, estratégia essencial para evitar interceptações e garantir a integridade das informações.

Embora inicialmente restrita a fins militares e acadêmicos, somente a partir de 1979 começou-se a estudar a possibilidade de sua aplicação em operações industriais, especialmente na interação com fornecedores e clientes. Em 1989, experimentos foram realizados no Canadá (Chevrolet–Pontiac–GM), Coreia do Sul e Japão, testando-se o envio de pedidos industriais por meio da internet. Entretanto, os altos custos de implantação e manutenção, aliados à limitada capacidade dos equipamentos disponíveis, impediram um avanço significativo naquele período.

Com a evolução tecnológica ao longo das décadas seguintes — incluindo o barateamento de hardware e a ampliação das redes — tornou-se viável integrar processos internos de produção às operações de fornecedores, clientes e unidades remotas, ampliando a conectividade e permitindo novos modelos de gestão.

A partir da segunda metade do século XX, observou-se uma difusão acelerada dos sistemas de controle, inicialmente mecânicos, depois eletroeletrônicos, chegando aos controles analógicos e, posteriormente, aos sistemas digitais e computacionais. No início do século XXI, a massificação da internet e a ampla disponibilidade de computadores, tablets e smartphones criaram o ambiente ideal para o surgimento do conceito de Indústria 4.0, que integra tecnologias digitais avançadas para elevar produtividade,

flexibilidade, qualidade e capacidade de gestão, além de fomentar novos modelos de negócios (Antunes, 2020).

Para que a Indústria 4.0 seja efetivamente implementada, certas estruturas tecnológicas são indispensáveis. (Sacomano et al., 2018) destacam, dentre outros elementos, a automação, a comunicação Máquina–Máquina (Machine to Machine – M2M) e a Inteligência Artificial (IA).

A automação refere-se à execução de tarefas sem intervenção humana direta, por meio de sistemas capazes de operar de maneira autônoma. Para a implantação da Indústria 4.0, é essencial que a linha de produção apresente, no mínimo, algum nível de automação. Contudo, não se exige que todo processo seja totalmente autônomo: muitos arranjos produtivos modernos combinam tecnologias automatizadas com trabalho humano, resultando em linhas híbridas altamente eficientes.

Um exemplo ilustrativo é a fabricação de notebooks. De acordo com De Almeida (Castro Melo, 2017), em uma empresa produtora de placas de circuito impresso, componentes pequenos — como resistores, capacitores, indutores e microchips — são instalados por robôs com elevada precisão e velocidade. Já etapas como montagem da carcaça, encaixe de módulos de memória e testes finais permanecem sob responsabilidade de operadores humanos. Assim, a automação é aplicada de forma estratégica, em tarefas que demandam alta precisão, repetitividade ou que representem risco ocupacional.

A Comunicação Máquina–Máquina (M2M) consiste na troca autônoma de informações entre dispositivos, por meio de conexões cabeadas ou sem fio (internet, intranet, Bluetooth, entre outras). Esse tipo de comunicação permite que máquinas interajam diretamente com sistemas computacionais ou servidores centrais.

Nesse contexto, a Inteligência Artificial torna-se fundamental. A IA utiliza métodos computacionais capazes de simular processos de tomada de decisão, identificar padrões, otimizar operações e contribuir para o aprimoramento contínuo dos sistemas produtivos. Por isso, constitui elemento chave para o desenvolvimento de soluções alinhadas aos princípios da Indústria 4.0 (Antunes, 2020).

A principal vantagem da adoção da Indústria 4.0 é a possibilidade de monitorar, acionar e intervir em processos produtivos de maneira remota, aumentando significativamente a confiabilidade operacional. Além disso, processos podem ser simulados previamente, reduzindo riscos, identificando gargalos e permitindo a tomada de decisão

baseada em dados.

Sistemas industriais baseados nesse paradigma são geralmente comandados por dispositivos eletrônicos conectados à internet ou a redes internas (intranet), possibilitando programação, controle, gerenciamento e cooperação em tempo real. Os benefícios não se limitam ao setor industrial tradicional: áreas como agricultura, saneamento (tratamento de água e esgoto), comércio eletrônico, varejo, restaurantes e diversos outros serviços também passaram a adotar tecnologias oriundas da Indústria 4.0 (Sacomano et al., 2018).

2.3 Aprendizagem de Máquina

A aprendizagem de máquina constitui um dos principais campos da Inteligência Artificial e tem como finalidade desenvolver métodos computacionais capazes de aprender a partir de dados, aprimorando seu desempenho conforme novas informações são incorporadas. Seu propósito central é permitir que sistemas computacionais adquiram conhecimento automaticamente, reduzindo a necessidade de intervenção humana.

De forma prática, um sistema de aprendizagem de máquina pode ser compreendido como um algoritmo capaz de realizar decisões ou previsões com base em padrões extraídos de experiências anteriores, tais como exemplos, análises históricas ou soluções previamente bem-sucedidas (Monard; Baranauskas, 2003). Assim, o modelo aprende relações e regularidades presentes nos dados e utiliza esse aprendizado para prever novos valores ou classificar novas instâncias.

2.3.1 Regressão

Os primeiros métodos de regressão remontam aos trabalhos de Legendre (1805) e Gauss (1809). Embora concebida há mais de dois séculos, a regressão permanece como uma ferramenta fundamental em problemas de estimação de funções e análise de relações entre variáveis. Com a crescente disponibilidade de dados — tanto em quantidade quanto em complexidade —, técnicas de regressão tornaram-se ainda mais relevantes em aplicações modernas (Izbicki; Santos, 2020).

Em aprendizagem de máquina supervisionada, a regressão busca prever valores contínuos a partir de variáveis independentes. Neste trabalho, seu uso está associado à previsão da temperatura de saída do sistema em estudo.

2.3.1.1 Random forest

O Random Forest Regressor é um meta-estimador que constrói múltiplas árvores de decisão utilizando subamostras do conjunto de dados, combinando-as por meio da média de suas previsões. Esse processo reduz o sobreajuste (overfitting) e aumenta a acurácia geral do modelo.

Segundo Breiman (Breiman, 2001), durante o treinamento o algoritmo define a melhor divisão para amostras com valores faltantes (NaNs) considerando o ganho de informação. Durante a etapa de predição, a amostra é direcionada ao ramo correspondente ou, se não houver histórico de NaNs para aquela variável, ao nó que concentra a maior quantidade de observações. Isso confere robustez ao método mesmo em cenários com dados incompletos.

2.3.1.2 Multi-Layer Perception (MLP)

O Perceptron de Múltiplas Camadas (MLP) é uma rede neural artificial composta por camadas de neurônios interligados, cada qual empregando funções de ativação não lineares. Essa característica permite ao modelo aprender padrões complexos e relações não lineares entre as variáveis, sendo amplamente empregada em processamento de linguagem, reconhecimento de fala e análise de imagens. A aplicabilidade prática do MLP é evidenciada em Popescu (Popescu et al., 2009).

Por sua flexibilidade arquitetural e capacidade de atuar como aproximador universal de funções, o MLP ocupa papel central na aprendizagem profunda (deep learning).

2.3.1.3 Regressão Linear

A Regressão Linear é uma técnica estatística amplamente utilizada para modelagem e previsão, cujo objetivo é estimar o valor de uma variável dependente em função de uma ou mais variáveis independentes.

Conforme Matos (Matos, 1995), o "melhor ajuste" é obtido pelo Método dos Mínimos Quadrados, que minimiza a soma dos erros quadráticos entre os valores previstos e os valores observados. Assim, o modelo busca a linha (ou hiperplano) que melhor descreve a relação entre as variáveis

2.3.1.4 Regressão de Ridge

A Regressão Ridge (Regularização L2) aplica uma penalização aos coeficientes da regressão linear para reduzir a variância do modelo e mitigar o sobreajuste. Segundo Reynaldo (Reynaldo, 1997), esse método é especialmente eficaz em cenários de multicolinearidade, impedindo que parâmetros assumam valores excessivamente elevados e garantindo maior estabilidade numérica.

2.3.1.5 Regressão de Lasso

A Regressão Lasso (Regularização L1) é voltada ao aprimoramento do modelo e ao combate ao sobreajuste por meio da penalização do valor absoluto dos coeficientes. Seu hiperparâmetro de regularização, λ , controla o grau de penalidade aplicado. Em valores mais altos, coeficientes de variáveis pouco relevantes são reduzidos a zero, produzindo modelos parcimoniosos.

Esse mecanismo torna o Lasso especialmente útil em bases de dados de alta dimensionalidade, pois realiza seleção automática de variáveis. Como destacado por (Alcântara Junior, 2021), a capacidade de eliminar coeficientes diferencia o Lasso da regularização L2 (Ridge), que apenas reduz valores sem descartá-los.

2.3.1.6 SVR

O objetivo da SVR é encontrar uma função que aproxime a relação entre variáveis de entrada e a variável alvo contínua, minimizando o erro de forma controlada. Sua principal vantagem está na capacidade de modelar relações não lineares por meio de funções kernel, que projetam os dados para espaços de maior dimensão, onde a relação se torna linear.

Segundo Awad (Awad; Khanna, 2015), a SVR utiliza uma função de perda simétrica e define um "tubo" de tolerância ao redor da função estimada, dentro do qual erros menores que um limiar não são penalizados. Observações fora do tubo são penalizadas, permitindo ao modelo equilibrar complexidade e generalização.

2.3.1.7 Gradient Boosting Regressor

De acordo com Medina (Medina et al., 2024), o Gradient Boosting constrói modelos de conjunto (ensembles) por meio da combinação sequencial de modelos fracos — geralmente árvores de decisão. A cada iteração, o novo modelo corrige os erros residuais do modelo anterior, minimizando gradualmente a função de perda.

Esse procedimento resulta em modelos capazes de capturar relações altamente complexas e não lineares, com excelente desempenho preditivo.

2.3.1.8 XGBoost

O XGBoost (eXtreme Gradient Boosting) é uma implementação otimizada do Gradient Boosting, criada para maximizar eficiência computacional, velocidade e escalabilidade. O algoritmo incorpora regularização, expansão do gradiente para segunda ordem e processamento otimizado de dados esparsos.

O estudo de Chen e Guestrin (Chen; Guestrin, 2016) demonstra como essas melhorias tornam o XGBoost uma das técnicas mais eficientes e precisas para grandes volumes de dados, tanto em regressão quanto em classificação.

2.3.1.9 LightGBM

O LightGBM é um framework de Gradient Boosting projetado para alta eficiência, especialmente em conjuntos de dados extensos. Seu mecanismo de crescimento de árvore baseado em folhas (leaf-wise) reduz o custo computacional e acelera o treinamento, embora possa aumentar o risco de sobreajuste em bases pequenas.

Como destacado por Horbe (Hörbe; Erol, 2025), o LightGBM apresenta vantagens significativas no tempo de busca por hiperparâmetros, reduzindo custos computacionais e favorecendo aplicações industriais.

2.3.2 Lógica Fuzzy

A Lógica Fuzzy (ou lógica difusa) surgiu de forma embrionária em pesquisas sobre incerteza na década de 1920, mas sua aplicação prática consolidou-se apenas a partir dos anos 1980. Sua principal motivação reside no fato de que muitos processos reais não se limitam a valores binários (0 ou 1), como na lógica booleana convencional. À medida que sistemas se tornaram mais complexos, foi necessário adotar métodos capazes de representar graus intermediários de pertinência (Aguado; Cantanhede, 2010).

Para ilustrar, Marro (Marro et al., 2010) compara a lógica clássica e a lógica fuzzy no contexto da categorização de alturas humanas. Na lógica tradicional, uma pessoa pertence exclusivamente a uma categoria (baixa, média ou alta). Já na lógica fuzzy, cada pessoa recebe um grau de pertinência — variando de 0 a 1 — indicando o quanto pertence a cada categoria. Assim, valores intermediários permitem uma modelagem mais próxima da realidade.

2.3.3 Metodos avaliativos

2.3.3.1 Coeficiente de determinação (R^2)

Um dos Métodos avaliativos do projeto é comprovada a partir da função do coeficiente de determinação (R^2), que expressa a quantidade de variância dos dados que é explicada pelo modelo linear, assim quanto maior o R^2 mais ele se aproxima da amostra, podendo sempre variar de 0 a 1, apresentando assim uma resposta em porcentagem em relação a aproximação das retas.

O modelo utilizado no coeficiente de determinação é representado por:

Soma dos quadrados totais:

$$SQ_{tot} = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (2)$$

Onde:

SQ_{tot} - Soma dos quadrados totais;

n - Número de observações;

y_i - Valor observado;

\bar{y} - Média geral do modelo.

Soma dos quadrados dos resíduos:

$$SQ_{res} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

Onde:

SQ_{res} - Soma dos quadrados dos resíduos;

n - Número de previsões;

y_i - Valor observado;

\hat{y}_i - Valor previsto.

Formula final do R^2 :

$$R^2 = 1 - \frac{SQ_{res}}{SQ_{tot}} \quad (4)$$

Onde:

R^2 - Coeficiente de determinação;

SQ_{res} - Soma dos quadrados dos resíduos;

SQ_{tot} - Soma total dos quadrados.

Quanto menor o valor dos resíduos de zero, tende a zero a fração das somas dos quadrados apresentando um modelo linear cada vez mais próximo do sistema(Chiode, 2021).

2.3.3.2 Erro Quadrático Médio (MSE)

Outro método utilizado para a análise de resultados do projeto foi o erro quadrático médio (MSE), que calcula a média das diferenças quadráticas entre os valores previstos pelo modelo e os valores reais, como é apresentada na formula abaixo:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

Onde:

MSE - Erro quadrático médio;

n - Número de pontos de dados;

y_i - Valores observados;

\hat{y}_i - Valores previstos.

Como demonstra (Alves; Gomes, 2020), quanto maior o erro, mais prejudicado serão os resultados devido a equação elevar ao quadrado o erro obtido. Por sua vez quanto mais próximo de zero melhor o ajuste entre o valor observado e o previsto.

2.3.3.3 Raiz do Erro Quadrático Médio (RMSE)

A Raiz do Erro Quadrático Médio (RMSE) é uma métrica utilizada para interpretação dos resultados obtidos pelo Erro Quadrático Médio, visto que ela retorna a variável alvo para a unidade de origem, como é representado abaixo:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (6)$$

Onde:

$RMSE$ - Raiz do erro quadrático médio;

n - Número de pontos de dados;

y_i - Valores observados;

\hat{y}_i - Valores previstos.

A Raiz do Erro Quadrático Médio se mostra mais eficaz quando o quesito é calcular a sensibilidade do erro do modelo ou em aplicações de assimilação de dados, devido a sua co-relação ao Erro Quadrático Médio e penalizar de forma mais rigorosa os erros graves obtidos(Chai; Draxler et al., 2014).

2.3.3.4 Erro Absoluto Médio (MAE)

É descrita como uma métrica de avaliação complementar aos critérios tradicionais baseados em quadrados (como MSE e RMSE). A MAE mede, em média, o quanto as previsões se afastam dos valores reais, considerando a magnitude dos erros de forma linear e não quadrática(Montgomery; Peck; Vining, 2021).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - p_i| \quad (1)$$

Onde:

MAE - Erro médio absoluto;

n - Número total de pontos de dados;

y_i - Valor observado;

p_i - Valor previsto pelo modelo.

2.3.3.5 Gráfico de Resíduos:

O Gráfico de Resíduos é uma ferramenta gráfica essencial na análise de regressão utilizada para avaliar a adequação e a validade de um modelo estatístico. Ele plota os resíduos (erros de previsão) do modelo em função de alguma outra variável, geralmente os valores previstos (\hat{Y}_i) ou as variáveis explicativas (X_i).

Conforme (Draper, 1998), utiliza-se a análise de gráfico de resíduos em no seguintes quesitos:

- **Verificar a não-normalidade.**
- **Verificar efeitos temporais se a ordem dos dados for conhecida.**
- **Verificar a variância não constante e a possível necessidade de uma transformação em Y .**
- **Verificar a curvatura de ordem superior do que a ajustada no X**

O Gráfico de Resíduos é o que garante a independência dos erros. Para que o modelo seja válido, os erros de previsão de uma observação não devem estar relacionados aos erros de outras observações, especialmente aquelas coletadas antes ou depois. O padrão visual ideal, que confirma o pressuposto de independência, é uma nuvem de pontos totalmente aleatória, sem padrões ou agrupamentos, dispersa uniformemente acima e abaixo da linha central ($y = 0$).

2.4 Trabalhos Correlatos

O controle de temperatura desempenha papel central em processos industriais, sendo o controlador Proporcional–Integral–Derivativo (PID) a estratégia mais amplamente utilizada. Em sua forma clássica, os ganhos proporcional, integral e derivativo são ajustados por técnicas específicas, dentre as quais se destacam os métodos de Ziegler–Nichols, Lugar das Raízes, otimização do ganho integral com restrição M, Kadda–Tau e controle por modelo interno. Entretanto, muitos processos industriais

apresentam elevada complexidade, com forte não linearidade, incertezas dinâmicas e defasagens temporais, o que exige reajustes frequentes sempre que o modelo ou as condições operacionais se alteram. Nesses cenários, embora os controladores PID sejam consagrados e confiáveis, sua eficiência tende a diminuir, tornando-se limitados em aplicações industriais modernas. Como alternativa, métodos baseados em lógica fuzzy vêm sendo incorporados ao ajuste de parâmetros, permitindo uma sintonia automática mais rápida e precisa.

Em (Gao; Zhu, 2025), foi analisado um sistema de aquecimento aplicado à distribuição de cola industrial, composto por um aquecedor tiristor multicanal (TH), um dispensador de cola, um dispositivo multicanal de amostragem de temperatura e um controlador integrado. O TH, as placas cerâmicas com coeficiente de temperatura positivo (PTC), o termopar e o dispositivo em teste (DUT) foram instalados diretamente no ambiente produtivo. Durante a operação, o TH fornece calor ao DUT por meio das placas cerâmicas, enquanto sensores independentes monitoram a temperatura e enviam dados ao controlador, que opera com cinco canais configurados para diferentes valores de referência. A partir dessa arquitetura, foi desenvolvido um sistema de controle em malha fechada responsável pela geração de sinais PWM para comutação do tiristor e modulação da potência aplicada aos elementos de aquecimento. A sintonia inicial do PID foi realizada pelo método de Ziegler–Nichols, e os experimentos compararam três abordagens: PID convencional, PID adaptativo e PID fuzzy. Os resultados indicaram que o método proposto reduziu o overshoot entre 2 °C e 4 °C, aumentou a velocidade de convergência em cerca de 30% e diminuiu a flutuação térmica após perturbações para aproximadamente 0,2 °C, com desempenho comparável ao controlador PID fuzzy (Gao; Zhu, 2025).

Outro trabalho relevante avaliou a implementação de um sistema de controle de temperatura em cultivo hidropônico de alface durante o verão na China (Cai et al., 2025). O estudo evidenciou que a temperatura tanto do ambiente quanto da solução nutritiva é determinante para o crescimento das plantas e para o rendimento agrícola. Inicialmente, foi projetado um controlador PID digital, que apresentou ganhos importantes em precisão e tempo de resposta; contudo, demandava retunings frequentes para manter o desempenho desejado. Métodos tradicionais baseados em fórmulas gerais mostraram-se pouco eficientes e com elevada exigência de tempo de calibração, revelando-se inadequados frente às variações ambientais complexas. Em seguida,

adotou-se um controlador PID fuzzy, capaz de lidar com incertezas e não linearidades por meio do ajuste automático de K_p , K_i e K_d . Posteriormente, foi desenvolvido um controlador PID DBO–Fuzzy adaptativo, integrando sensores de alta precisão e um sistema de bombeamento térmico como base de hardware. Nessa configuração, o algoritmo DBO passou a otimizar o controle em tempo real, permitindo simulações, testes de controle térmico e validação direta no cultivo. Ensaios de resposta ao degrau compararam PID convencional, PID fuzzy e PID DBO–Fuzzy, avaliando overshoot, tempo de estabilização e erro em regime permanente. O PID DBO–Fuzzy apresentou o melhor desempenho, com resposta mais rápida, menor sobrelevação e menor erro final. Em experimentos práticos, o sistema foi aplicado em três áreas idênticas, cada uma operando com um dos métodos. A temperatura da solução nutritiva, idealmente em 22 °C, foi monitorada ao longo do verão; o controlador DBO–Fuzzy manteve a temperatura mais próxima do valor ideal mesmo em condições desfavoráveis, enquanto o PID convencional apresentou maiores desvios térmicos. O estudo confirmou ainda a influência direta da temperatura da solução nutritiva no crescimento das plantas, reforçando a importância do controle térmico no cultivo hidropônico (Cai et al., 2025).

Em (Li; Wang, 2024), os autores propõem uma solução para aprimorar o controle inteligente de temperatura em instrumentação industrial, destacando as limitações de desempenho de PIDs tradicionais em ambientes sujeitos a variações constantes e necessidade de retunings frequentes. Para superar essas restrições, é apresentado um controlador Fuzzy PID (FPID), capaz de ajustar automaticamente os coeficientes do controlador em tempo real por meio de inferência baseada em lógica fuzzy. A fim de elevar a eficiência do controle, o estudo aplica o algoritmo Multi-Strategy Grey Wolf Optimization (MGWO), uma versão aprimorada da técnica Grey Wolf Optimization (GWO), para o ajuste automático dos parâmetros do controlador fuzzy. A integração resulta no modelo MGWO–FPID, avaliado experimentalmente em um instrumento térmico industrial real, com simulações realizadas em MATLAB. O desempenho foi comparado a dois modelos avançados: PSO–FPID (otimizado por Particle Swarm Optimization) e ISOA–FPID (baseado no Improved Seeker Optimization Algorithm). Os modelos foram avaliados por métricas robustas de desempenho, como tempo de estabilização, overshoot, erro médio absoluto (MAE), precisão, recall, F1-score e área sob a curva (AUC). O MGWO–FPID obteve o melhor desempenho entre os métodos, com ausência de overshoot, tempo de estabilização de 2,08 s, MAE de 8,53, precisão

de 99,52%, recall de 95,37%, F1-score de 96,14% e AUC de 0,995, evidenciando maior eficiência, robustez e confiabilidade para aplicações industriais que exigem controle térmico avançado (Li; Wang, 2024).

Alexandre (Alexandre et al., 2016) realizou uma comparação entre três estratégias de controle em uma planta industrial de nível e temperatura (tanque de aquecimento): PID convencional, PID Fuzzy e controle preditivo (Model Predictive Control – MPC). A motivação do estudo decorre da crescente necessidade de controle eficaz em processos multivariáveis e não lineares, típicos da indústria. As três arquiteturas foram implementadas e simuladas em malha fechada no ambiente MATLAB/Simulink, sob condições operacionais idênticas, avaliando-se tempo de resposta, estabilidade, overshoot e capacidade de acomodação a variações operacionais, tanto na malha de nível quanto na de temperatura. Os resultados mostraram que o PID Fuzzy apresentou o melhor desempenho dinâmico, com tempo de resposta de aproximadamente 30 s na malha de nível e 36 s na malha de temperatura, em comparação com 125 s e 27 s do PID convencional, respectivamente, e 76 s e 80 s do controlador preditivo. Concluiu-se que, embora o MPC supere o PID tradicional em cenários com acoplamento e não linearidade, o PID Fuzzy mostrou-se mais robusto e eficaz para a planta em estudo (Alexandre et al., 2016).

Por fim, em (Junior et al., 2020) foi apresentada a implementação de um controlador de temperatura baseado em lógica fuzzy em uma planta experimental didática, como alternativa aos controladores clássicos que exigem modelagem matemática precisa. O controlador fuzzy do tipo Mamdani foi desenvolvido em LabVIEW e integrado a um Arduino UNO para aquisição de dados. Foram definidas variáveis linguísticas para setpoint, temperatura medida e saída (PWM), construídas funções de pertinência triangulares e trapezoidais e estabelecidas 15 regras de inferência do tipo Se–Então. Ensaios foram realizados para diferentes setpoints (35 °C, 38 °C e 41 °C), avaliando o desempenho do controle. Os resultados indicaram que o controlador foi capaz de estabilizar a temperatura com erros inferiores a 3%, apresentando erros finais de 0,57%, 0,80% e 2,20%, respectivamente. Apesar de algumas limitações em faixas intermediárias, nas quais determinadas regras não são ativadas, o estudo demonstrou a viabilidade e o bom desempenho de controladores fuzzy em aplicações práticas, especialmente quando se busca flexibilidade, simplicidade de implementação e boa resposta em sistemas não lineares (Junior et al., 2020).

3 MATERIAL E MÉTODOS

Este capítulo descreve a base de dados utilizada e a metodologia adotada para o desenvolvimento, treinamento e avaliação dos modelos de aprendizagem de máquina propostos neste trabalho.

3.1 Base de dados

A base de dados empregada neste estudo é de domínio público e foi obtida na plataforma Kaggle, proveniente de um sistema de manufatura inteligente voltado ao controle de temperatura. Os valores foram registrados automaticamente pelo sistema a cada 5 minutos, totalizando 1000 amostras de operação, o que permite analisar o comportamento dinâmico do processo sob diferentes condições.

A Figura 1 apresenta um recorte ilustrativo da base de dados, evidenciando a estrutura tabular das medições, com múltiplos parâmetros de processo e variáveis de controle registradas em cada instante de amostragem.

	Timestamp	Current Temperature (°C)	Setpoint Temperature (°C)	Temperature Error (°C)	PID Control Output (%)	Fuzzy PID Control Output (%)	Overshoot (°C)	Response Time (s)	Steady-State Error (°C)	Ambient Temperature (°C)	Humidity (%)	PID Kp	PID Ki	PID Kd	Fuzzy Rule Base Parameters
0	2025-01-31 12:09:12.524887	368.727006	400	31.272994	67.405317	65.022374	6.727030	47.159876	0.787271	23.944771	50.387995	1.220268	0.465431	0.249456	2
1	2025-01-31 12:14:12.524887	397.535715	400	2.464285	81.676038	79.145826	7.966814	54.162970	0.946871	22.517159	51.867725	1.187283	0.310144	0.233165	0
2	2025-01-31 12:19:12.524887	386.599697	400	13.400303	94.917833	98.980379	2.504679	52.804828	1.709095	24.617184	58.312458	0.595754	0.389964	0.170462	0
3	2025-01-31 12:24:12.524887	379.932924	400	20.067076	89.288995	86.784457	6.248741	34.616997	0.680009	23.839349	57.667684	1.422572	0.274419	0.342907	1
4	2025-01-31 12:29:12.524887	357.800932	400	42.199068	92.262446	89.981943	5.717460	34.477484	1.739299	22.471612	53.506427	1.068472	0.352014	0.290650	2
...
995	2025-02-03 23:04:12.528890	354.579104	400	45.420896	86.278206	89.931164	4.421070	56.921920	1.368871	23.739211	52.731596	1.014929	0.443062	0.243778	0
996	2025-02-03 23:09:12.528890	395.865679	400	4.134321	98.264585	94.837317	3.344012	33.581430	1.006441	24.001111	52.095192	0.665370	0.459004	0.470683	2
997	2025-02-03 23:14:12.528890	356.840932	400	43.159068	62.758321	60.856199	3.945723	39.835285	1.530298	24.734731	54.553287	1.477713	0.478683	0.379463	2
998	2025-02-03 23:19:12.528890	397.511868	400	2.488132	62.282189	60.182644	5.299406	54.472361	0.970581	23.036861	59.084375	0.607925	0.258995	0.202871	0
999	2025-02-03 23:24:12.528890	372.300289	400	27.699711	71.287483	75.001623	1.613674	47.919371	0.298763	23.864522	51.044803	1.431944	0.186856	0.325634	2

Figura 1 – Amostra da base de dados utilizada.

A escolha desta base de dados justifica-se por três motivos principais: (i) trata-se de um cenário realista de controle térmico em ambiente industrial; (ii) contém simultaneamente sinais de controladores PID e Fuzzy-PID, permitindo comparar abordagens e treinar modelos de regressão; e (iii) apresenta quantidade suficiente de amostras para avaliação estatisticamente consistente dos algoritmos de aprendizagem de máquina.

3.1.1 Parâmetros de entrada

Os parâmetros utilizados como variáveis independentes (features de entrada) foram:

- **Erro de Temperatura (Temperature Error) (°C):** diferença instantânea entre a temperatura de referência e a temperatura atual do sistema.
- **Saída do Controle PID (PID Control Output) (%):** sinal de controle gerado pelo algoritmo PID tradicional, aplicado ao atuador do sistema térmico.
- **Sobressinal (Overshoot) (°C):** valor máximo de temperatura que ultrapassa a referência antes da estabilização do sistema.
- **Tempo de Resposta (Response Time) (s):** tempo necessário para que a resposta do sistema se aproxime da referência dentro de uma faixa aceitável.
- **Erro em Regime Permanente (Steady-State Error) (°C):** diferença final entre a temperatura de referência e a temperatura real após o sistema atingir regime estacionário.
- **Temperatura Ambiente (Ambient Temperature) (°C):** temperatura medida no entorno do sistema, que influencia a troca de calor com o meio.
- **Umidade (Humidity) (%):** nível de umidade relativa do ar, que também pode afetar a dinâmica térmica em determinados ambientes industriais.
- **PID K_p :** ganho proporcional do controlador PID, associado à correção imediata proporcional ao erro atual.
- **PID K_d :** ganho derivativo, responsável por atuar sobre a taxa de variação do erro, contribuindo para amortecer oscilações.
- **PID K_i :** ganho integral, relacionado à correção do erro acumulado ao longo do tempo, reduzindo o erro em regime permanente.
- **Parâmetros da Base de Regras Fuzzy (Fuzzy Rule Base Parameters):** parâmetros que descrevem a lógica utilizada pelo controlador Fuzzy (regras de inferência, funções de pertinência, entre outros), representados numericamente na base de dados.

Em conjunto, esses parâmetros descrevem tanto o estado do processo térmico quanto a configuração do controlador PID/Fuzzy utilizado, fornecendo um espaço de atributos rico para a construção de modelos de regressão.

3.1.2 Parâmetro de saída

O parâmetro utilizado como variável dependente (saída alvo) foi:

- **Saída do Controle PID Fuzzy (Fuzzy PID Control Output) (%)**: sinal de controle gerado pela versão aprimorada do controlador, que integra lógica Fuzzy ao PID tradicional.

Dessa forma, o problema de aprendizagem de máquina é formulado como uma tarefa de regressão supervisionada, na qual se busca aproximar a função que mapeia os parâmetros de processo e de controle para a saída do controlador Fuzzy-PID. Em termos práticos, o modelo de ML aprende a reproduzir (ou aproximar) a decisão do controlador inteligente a partir das condições de operação.

3.2 Método aplicado

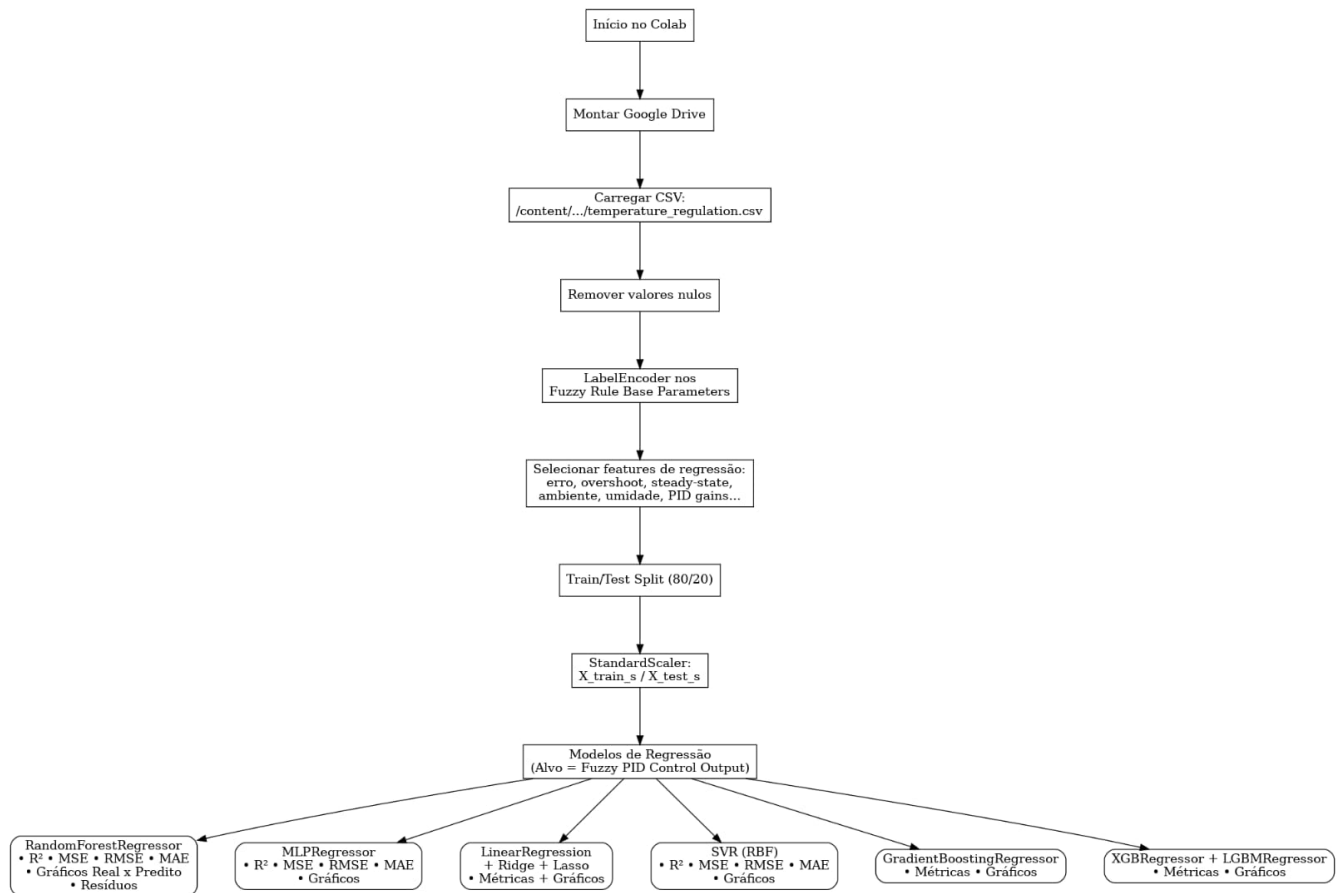


Figura 2 – Fluxograma da metodologia de Aprendizagem de Máquina Supervisionada

A metodologia representada no fluxograma da Figura 2 apresenta o pipeline adotado para o desenvolvimento do estudo, seguindo um fluxo típico de Aprendizagem de Máquina Supervisionada para problemas de regressão. As etapas abrangem desde a preparação dos dados até a avaliação comparativa dos modelos.

3.2.1 Preparação e carregamento de dados

- **Ambiente de desenvolvimento (Colab):** todo o processamento foi realizado no ambiente Google Colab, o que garante reprodutibilidade, facilidade de compartilhamento do código e acesso direto ao Google Drive.
- **Carregamento do arquivo CSV:** a base de dados descrita na Seção 3.1 foi armazenada em formato .csv em um diretório do Google Drive, sendo carregada no Colab por meio de bibliotecas como *pandas*. Nesta etapa, são conferidas dimensões, tipos de dados e consistência básica do conjunto.

Essa etapa inicial assegura que o dataset esteja acessível, íntegro e pronto para ser submetido às fases de pré-processamento e modelagem.

3.2.2 Pré-processamento e engenharia de *features*

- **Tratamento de valores nulos:** foi realizada uma limpeza inicial do conjunto de dados, identificando e tratando eventuais observações com valores ausentes. Dependendo do caso, registros foram removidos ou, quando justificável, aplicadas estratégias simples de imputação. Esse passo evita que modelos sejam enviesados por dados incompletos.
- **Codificação de variáveis não numéricas:** em situações em que variáveis categóricas estivessem presentes, seria empregada a técnica *LabelEncoder* para converter rótulos em representações numéricas, requisito para a maior parte dos algoritmos de ML. No conjunto específico utilizado, os atributos relevantes para a regressão encontram-se predominantemente em formato numérico.
- **Seleção de *features* para regressão:** as colunas que compõem o vetor de entrada foram selecionadas a partir dos parâmetros de processo e de controle descritos na Seção 3.1.1. A seleção levou em conta tanto a relevância física

das variáveis (influência direta no controle térmico) quanto testes exploratórios preliminares. Dessa forma, foram mantidos os atributos com maior potencial explicativo para a saída do controlador Fuzzy-PID.

A etapa de pré-processamento é fundamental para garantir que os modelos recebam dados consistentes, escalonados e representativos do fenômeno em estudo, reduzindo ruídos e possíveis fontes de viés.

3.2.3 Treinamento e avaliação inicial dos modelos

- **Divisão treino–teste:** o conjunto de dados foi dividido em duas partes: 80% para treino e 20% para teste. O subconjunto de treino é utilizado para ajuste dos parâmetros dos modelos, enquanto o subconjunto de teste é reservado para avaliação de desempenho em dados nunca vistos, permitindo estimar a capacidade de generalização.
- **Padronização (*StandardScaler*):** tanto o conjunto de treino quanto o de teste foram submetidos à padronização, de modo que as *features* passem a ter média zero e desvio-padrão unitário. Esse procedimento é particularmente importante para modelos sensíveis à escala dos dados, como SVR, MLP e alguns métodos baseados em gradiente, prevenindo que atributos de maior magnitude dominem a função de custo.

Após essa etapa, os dados encontram-se em condições adequadas para a fase de modelagem, garantindo comparabilidade entre diferentes algoritmos.

3.2.4 Modelagem e otimização

Nesta etapa, foi conduzido o teste de múltiplos modelos de regressão supervisionada, com o objetivo de comparar o desempenho e identificar o melhor preditor para a variável alvo *Fuzzy PID Control Output*. Para cada modelo, foram realizados ajustes de hiperparâmetros de forma sistemática, de modo a obter uma configuração com bom compromisso entre desempenho e complexidade.

Os modelos considerados foram:

- **RandomForestRegressor:** modelo baseado em conjuntos de árvores de decisão, no qual foram ajustados, principalmente, o número de estimadores

(*n_estimators*) e a profundidade máxima das árvores (*max_depth*). Esses parâmetros influenciam diretamente a capacidade do modelo de capturar relações complexas sem incorrer em sobreajuste excessivo.

- **MLPRegressor (Rede Neural MLP)**: modelo de rede neural de múltiplas camadas, com otimização do número de neurônios e camadas ocultas (*hidden_layer_sizes*), da função de ativação e da taxa de aprendizado (*learning rate*). Tais ajustes visam equilibrar capacidade de representação e estabilidade do treinamento.
- **Linear Regression, Ridge, Lasso**: aplicação de modelos de regressão linear simples e regularizada. No caso de Ridge e Lasso, o parâmetro de regularização (α) foi ajustado de forma a controlar o trade-off entre viés e variância, evitando tanto subajuste quanto sobreajuste.
- **SVR (kernel RBF)**: utilização de Máquina de Vetores de Suporte para regressão, com kernel de Função de Base Radial (RBF). Foram considerados parâmetros como C , ϵ e o parâmetro do kernel (γ), que controlam a largura da margem, a tolerância a erros e a flexibilidade do modelo.
- **GradientBoostingRegressor (GBR)**: modelo de *boosting* baseado em árvores de decisão, com otimização do número de estimadores, da taxa de aprendizado e da profundidade máxima. Esses hiperparâmetros determinam o ritmo de aprendizado e o grau de detalhamento das correções sucessivas.
- **XGBRegressor e LightGBMRegressor**: modelos de *gradient boosting* de alto desempenho, nos quais foram ajustados parâmetros como profundidade das árvores, número de estimadores e taxa de aprendizado. Esses algoritmos são projetados para lidar com grandes volumes de dados e interações complexas entre variáveis, mantendo boa eficiência computacional.

Embora o processo de otimização não tenha buscado a exaustão de todas as combinações possíveis, foram adotados critérios consistentes de ajuste (como buscas em grades reduzidas e experimentação iterativa), suficientes para estabelecer uma comparação justa entre as diferentes famílias de modelos.

3.2.5 Avaliação de desempenho

O desempenho de todos os modelos testados foi avaliado por meio de métricas clássicas de regressão e de análises gráficas, permitindo tanto uma avaliação quantitativa quanto qualitativa dos resultados:

- **Coeficiente de Determinação (R^2):** mede a proporção da variância da variável dependente explicada pelo modelo. Valores próximos de 1 indicam bom ajuste global.
- **Erro Quadrático Médio (MSE):** calcula a média dos quadrados dos erros de previsão, penalizando mais fortemente erros grandes. É útil para quantificar a magnitude dos desvios em relação aos valores reais.
- **Raiz do Erro Quadrático Médio (RMSE):** é a raiz quadrada do MSE, mantendo a mesma unidade da variável alvo e facilitando a interpretação prática dos erros.
- **Erro Absoluto Médio (MAE):** corresponde à média dos valores absolutos dos erros, fornecendo uma medida robusta da discrepância média entre valores observados e previstos.
- **Gráficos de resíduos:** permitem inspecionar visualmente a distribuição dos erros em função das previsões ou do alvo, auxiliando na verificação de homocedasticidade, ausência de padrões sistemáticos e identificação de possíveis outliers.
- **Gráficos de saída real vs. saída predita:** ilustram o grau de aderência entre as previsões dos modelos e os valores reais da variável alvo, sendo especialmente úteis para comparar qualitativamente o desempenho entre algoritmos distintos.

A combinação dessas métricas e análises gráficas garante uma avaliação abrangente dos modelos, permitindo não apenas identificar o melhor desempenho numérico, mas também compreender a adequação de cada abordagem ao problema de controle considerado.

4 RESULTADOS E DISCUSSÃO

A partir da aplicação apresentada na seção anterior foram obtidos diversos resultados, os quais foram organizados, comparados e analisados com o objetivo de identificar o modelo de aprendizagem de máquina com melhor desempenho para o problema de predição estudado. Nesta seção são discutidos os indicadores estatísticos calculados (R^2 , MSE, RMSE e MAE), bem como os gráficos de resíduos e de saída real versus saída predita, buscando relacionar os achados com as características dos modelos e com o contexto de aplicação.

4.1 Coeficiente de Determinação (R^2)

A Tabela 1 apresenta os resultados obtidos para o coeficiente de determinação (R^2), métrica que indica a proporção da variância da variável alvo explicada pelo modelo.

Tabela 1 – Resultados dos modelos de Machine Learning (ML) em termos de R^2 .

Machine Learning (ML)	R^2
Random Forest	0.9364
Multi-Layer Perception (MLP)	0.9087
Linear Regression	0.9406
Ridge Regression	0.9405
Lasso Regression	0.9407
SVR Regression	0.8623
GradientBoostingRegressor	0.9293
XGBoost	0.9356
LightGBM	0.9318

Observa-se que as variantes da regressão linear (Linear, Ridge e Lasso) apresentaram desempenhos praticamente idênticos e foram as líderes em poder explicativo, com $R^2 \approx 0,94$. Esse resultado sugere fortemente que a relação entre as variáveis de entrada e a variável alvo é predominantemente linear ou, ainda, que a complexidade adicional introduzida pelos modelos não lineares não resultou em ganho expressivo de capacidade preditiva para este conjunto de dados.

O fato de a Regressão Ridge e a Regressão Lasso apresentarem valores de R^2 praticamente iguais aos da Regressão Linear simples indica, também, que não há evidência de multicolinearidade severa ou de coeficientes instáveis que exigissem forte regularização. Ainda assim, a Lasso mantém a vantagem teórica de promover maior parcimônia, por favorecer modelos com menos variáveis relevantes.

Em contrapartida, o Support Vector Regression (SVR) obteve o menor R^2 (0,8623), indicando que explica uma fração menor da variabilidade dos dados em comparação aos demais modelos. Esse desempenho inferior pode estar associado à escolha do kernel, dos hiperparâmetros (como C , ϵ e parâmetros do kernel) ou à própria estrutura dos dados, que pode não favorecer a forma específica de generalização adotada pelo SVR.

4.2 Erro Quadrático Médio (MSE)

A Tabela 2 apresenta o desempenho dos modelos em termos de Erro Quadrático Médio (MSE), métrica que penaliza mais fortemente erros grandes.

Tabela 2 – Resultados dos modelos de Machine Learning (ML) em termos de MSE.

Machine Learning (ML)	MSE
Random Forest	9.3628
Multi-Layer Perception (MLP)	13.4456
Linear Regression	8.7494
Ridge Regression	8.7528
Lasso Regression	8.7327
SVR Regression	20.2767
GradientBoostingRegressor	10.4052
XGBoost	9.4838
LightGBM	10.0387

De forma consistente com os resultados de R^2 , as variantes da Regressão Linear (Lasso, Linear e Ridge) obtiveram os menores valores de MSE, sendo a Regressão Lasso o melhor modelo, com MSE igual a 8,7327. A proximidade entre os três valores reforça a hipótese de que o comportamento do sistema é bem descrito por uma relação linear, sem necessidade de modelos altamente complexos para capturar a dinâmica do problema.

Modelos de árvore e ensemble (Random Forest, Gradient Boosting, XGBoost e LightGBM) também apresentaram desempenho competitivo, com MSEs relativamente baixos e próximos entre si, o que indica boa capacidade de generalização. Contudo, não superaram as regressões lineares regulares, sugerindo que o ganho potencial de flexibilidade estrutural não compensou o aumento de complexidade.

Por outro lado, o SVR e o MLP apresentaram MSEs consideravelmente mais elevados, indicando maior sensibilidade a erros grandes. Isso pode refletir uma combinação de fatores, tais como: (i) configuração de hiperparâmetros não ideal; (ii) necessidade

de um conjunto de treinamento maior para esses modelos atingirem melhor ajuste; (iii) presença de ruídos ou padrões que não foram adequadamente capturados por suas arquiteturas.

4.3 Raiz do Erro Quadrático Médio (RMSE)

A Tabela 3 apresenta o desempenho em termos da Raiz do Erro Quadrático Médio (RMSE), que é a raiz quadrada do MSE e, portanto, mantém a mesma interpretação, porém na mesma unidade da variável alvo.

Tabela 3 – Resultados dos modelos de Machine Learning (ML) em termos de RMSE.

Machine Learning (ML)	RMSE
Random Forest	3.06
Multi-Layer Perception (MLP)	3.67
Linear Regression	2.96
Ridge Regression	2.96
Lasso Regression	2.96
SVR Regression	4.50
GradientBoostingRegressor	3.23
XGBoost	3.08
LightGBM	3.17

Como esperado, o comportamento do RMSE acompanha diretamente o do MSE. As regressões lineares apresentam RMSE em torno de 2,96 unidades da variável alvo, o que representa, para o contexto estudado, um erro médio de ordem moderada. Em sistemas térmicos ou de processo, por exemplo, esse valor pode ser interpretado em termos de graus de temperatura, permitindo avaliar se o erro é aceitável ou não para fins de controle.

O SVR e o MLP mantêm RMSEs mais elevados, reforçando a evidência de que esses modelos não se ajustaram tão bem ao conjunto de dados quanto os modelos lineares e de ensemble.

4.4 Erro Absoluto Médio (MAE)

A Tabela 4 apresenta o Erro Absoluto Médio (MAE), que mede o erro médio em valor absoluto, sem penalizar erros grandes de forma quadrática.

As três variantes da Regressão Linear (Linear, Ridge e Lasso) compartilham o menor MAE (2,55), indicando que, em média, os erros cometidos por esses modelos

Tabela 4 – Resultados dos modelos de Machine Learning (ML) em termos de MAE.

Machine Learning (ML)	MAE
Random Forest	2.60
Multi-Layer Perception (MLP)	3.03
Linear Regression	2.55
Ridge Regression	2.55
Lasso Regression	2.55
SVR Regression	3.61
GradientBoostingRegressor	2.77
XGBoost	2.61
LightGBM	2.67

são da ordem de 2,5 unidades da variável alvo. Esse resultado é coerente com os valores de MSE e RMSE e reforça a robustez dos modelos lineares neste problema.

A diferença entre o MAE do SVR (3,61) e do MLP (3,03) em relação ao MAE das regressões indica que esses modelos erram com maior frequência e por margens maiores. Isso se alinha com o comportamento observado em MSE/RMSE e sugere que, do ponto de vista prático, SVR e MLP tendem a produzir previsões menos confiáveis para o sistema analisado.

4.5 Gráficos de resíduos

Os gráficos a seguir apresentam a distribuição dos resíduos (diferença entre saída real e predita) para cada modelo:

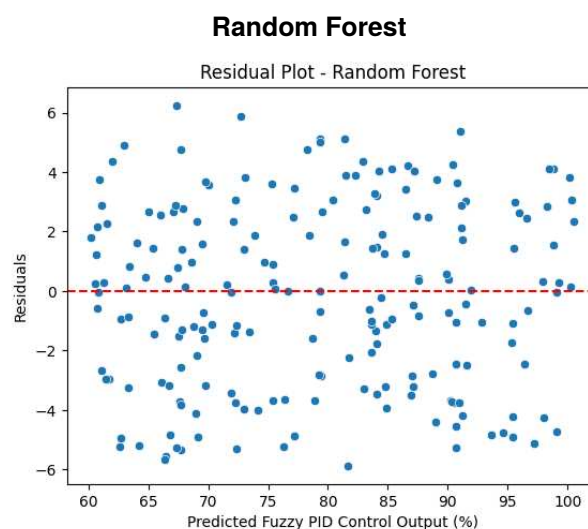


Figura 3 – Gráfico de resíduos do modelo Random Forest.

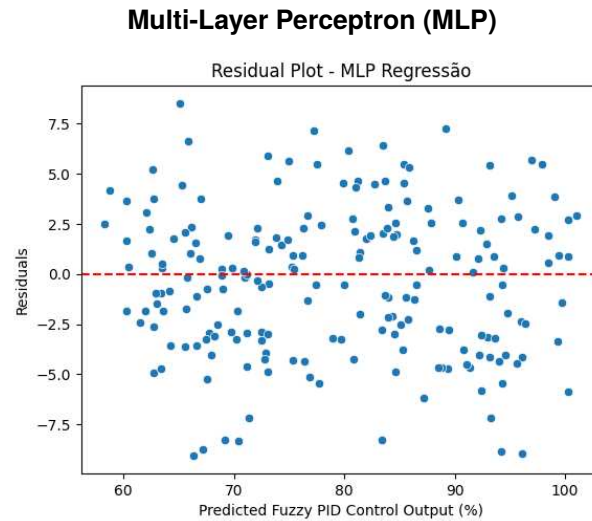


Figura 4 – Gráfico de resíduos do modelo MLP.

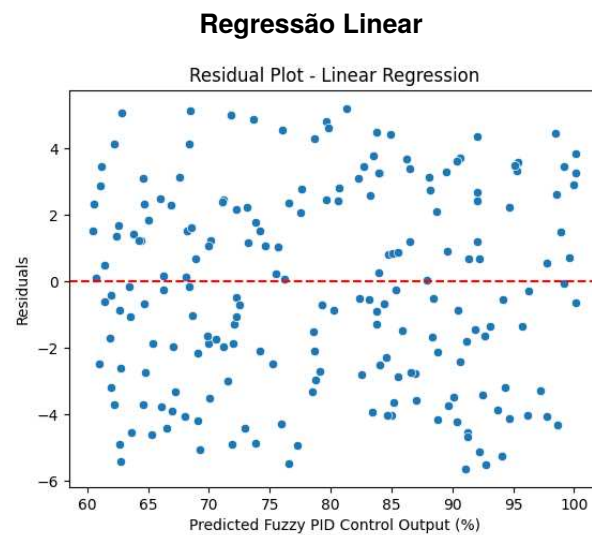


Figura 5 – Gráfico de resíduos do modelo Regressão Linear.

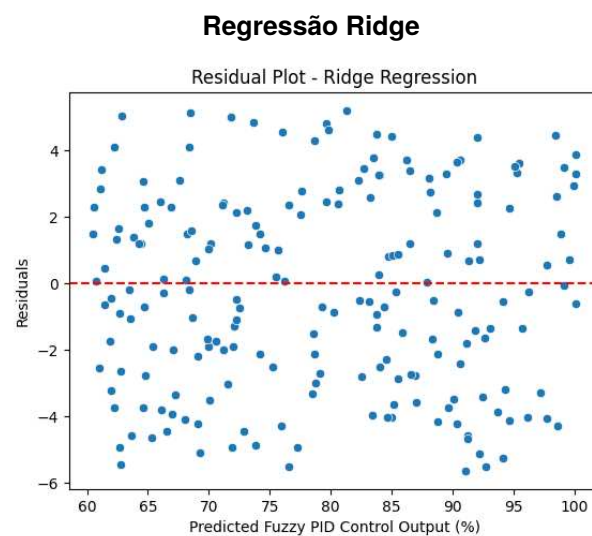


Figura 6 – Gráfico de resíduos do modelo Regressão Ridge.

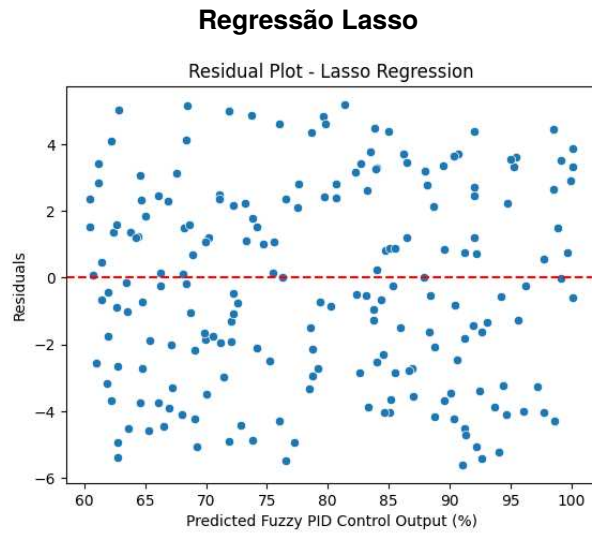


Figura 7 – Gráfico de resíduos do modelo Regressão Lasso.

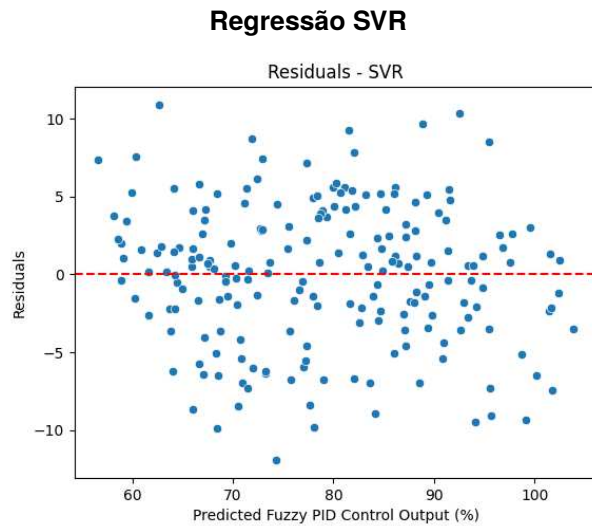


Figura 8 – Gráfico de resíduos do modelo SVR.

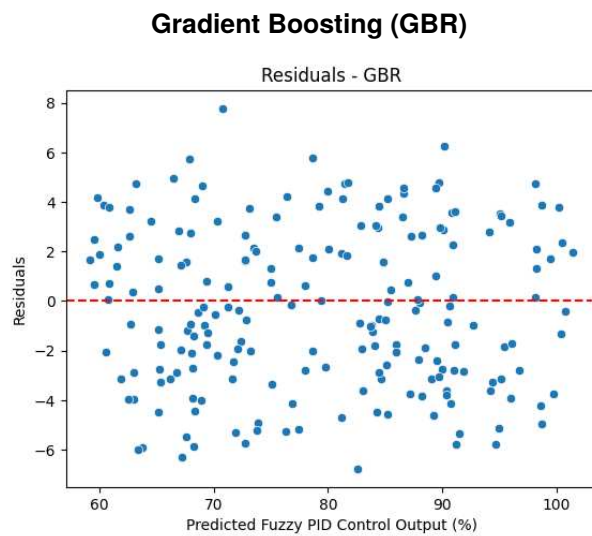


Figura 9 – Gráfico de resíduos do modelo Gradient Boosting.

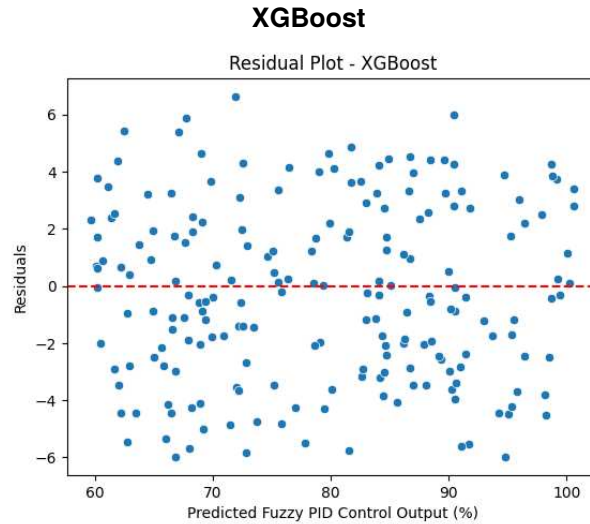


Figura 10 – Gráfico de resíduos do modelo XGBoost.

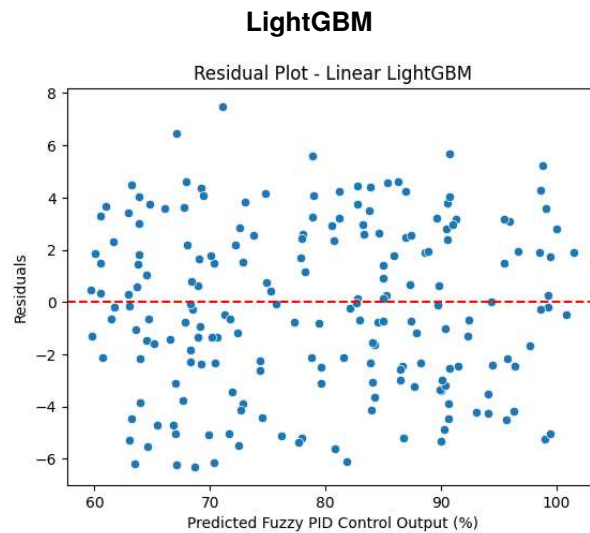


Figura 11 – Gráfico de resíduos do modelo LightGBM.

A análise visual dos resíduos indica, em todos os modelos, ausência de padrões estruturados evidentes, com distribuição aproximadamente aleatória em torno de zero. Esse comportamento sugere que não há violações graves das hipóteses clássicas de regressão (como tendência não capturada ou forte heterocedasticidade), principalmente nos modelos lineares e de ensemble.

Entretanto, observa-se que a dispersão dos resíduos varia entre os modelos. Nos regressores lineares, a maioria dos resíduos concentra-se em um intervalo relativamente estreito, com valores em torno de ± 5 unidades. Já no SVR, nota-se uma dispersão bem maior, com resíduos que ultrapassam 10 unidades em valor absoluto, o que está em consonância com o MSE elevado observado anteriormente. Esse comportamento reforça a conclusão de que o SVR não conseguiu capturar de forma adequada

a estrutura dos dados neste problema específico.

4.6 Gráficos de saída real versus saída predita

Os gráficos a seguir apresentam a comparação entre as saídas observadas e as saídas previstas para cada modelo:

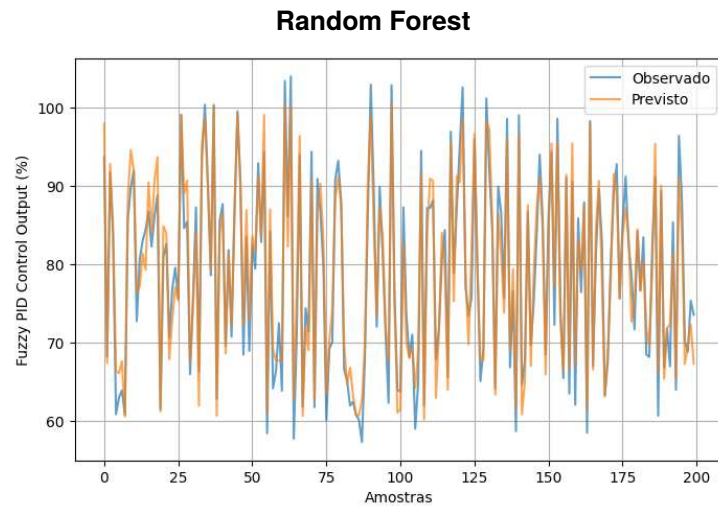


Figura 12 – Gráfico de saída real x prevista para o modelo Random Forest.

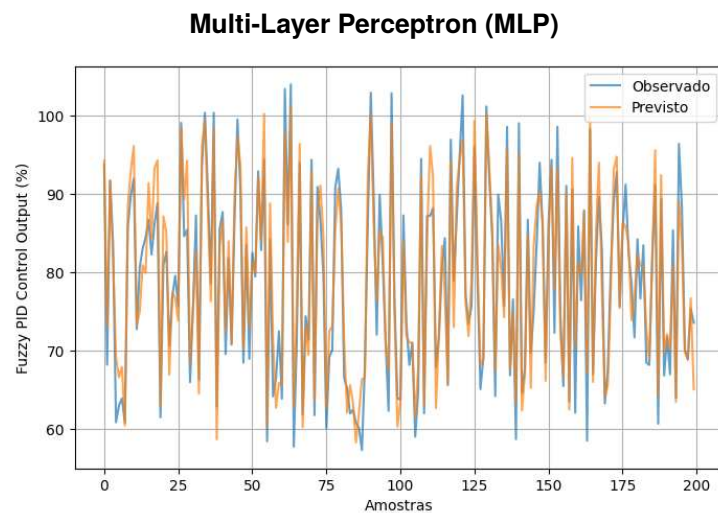


Figura 13 – Gráfico de saída real x prevista para o modelo MLP.

Regressão Linear

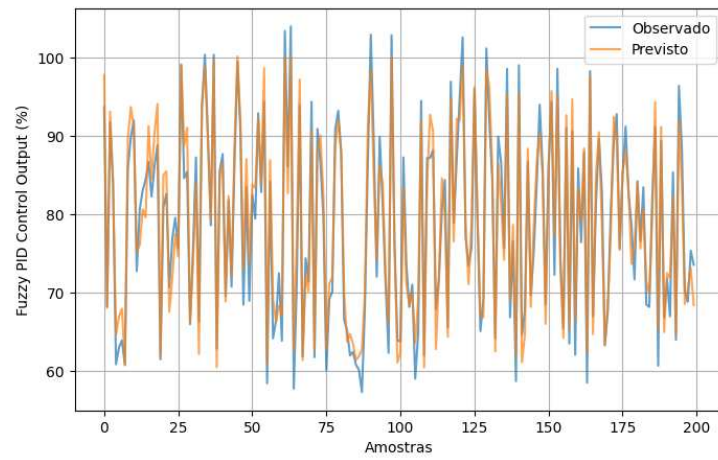


Figura 14 – Gráfico de saída real x prevista para o modelo Regressão Linear.

Regressão Ridge

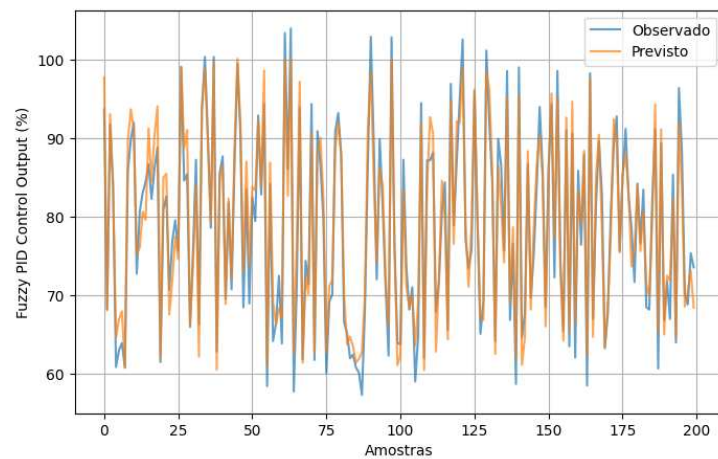


Figura 15 – Gráfico de saída real x prevista para o modelo Regressão Ridge.

Regressão Lasso

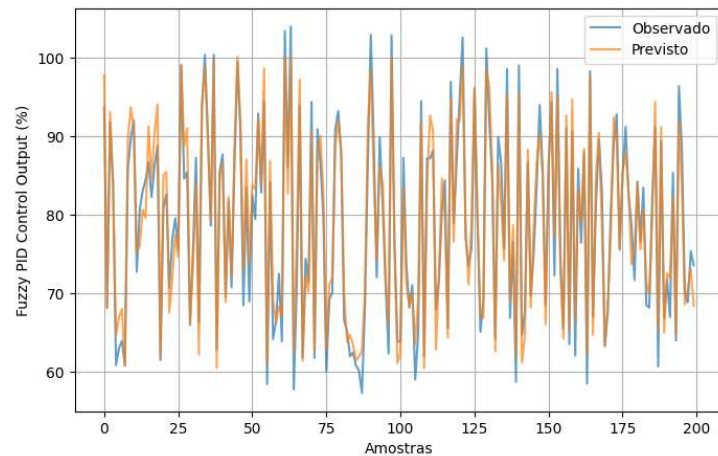


Figura 16 – Gráfico de saída real x prevista para o modelo Regressão Lasso.

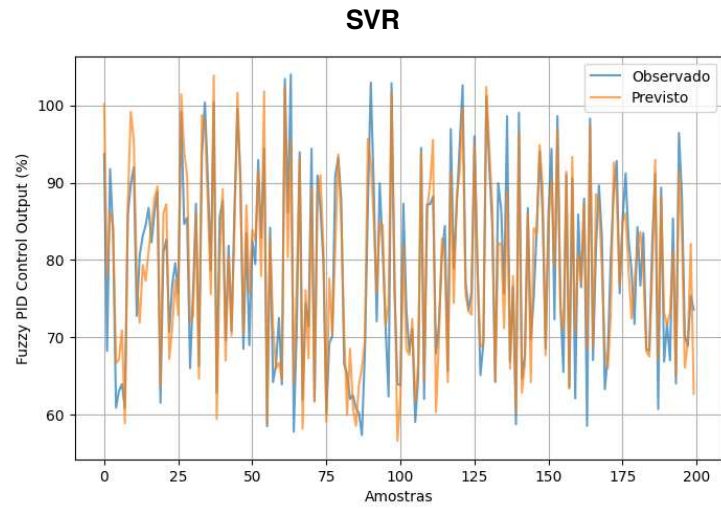


Figura 17 – Gráfico de saída real x prevista para o modelo SVR.

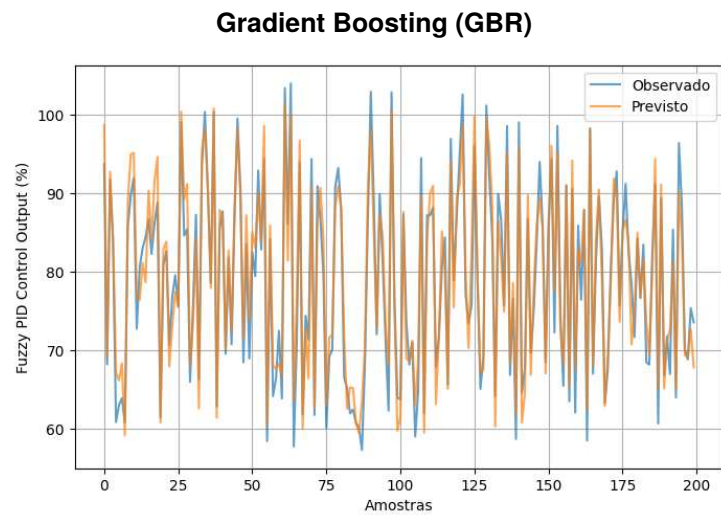


Figura 18 – Gráfico de saída real x prevista para o modelo Gradient Boosting.

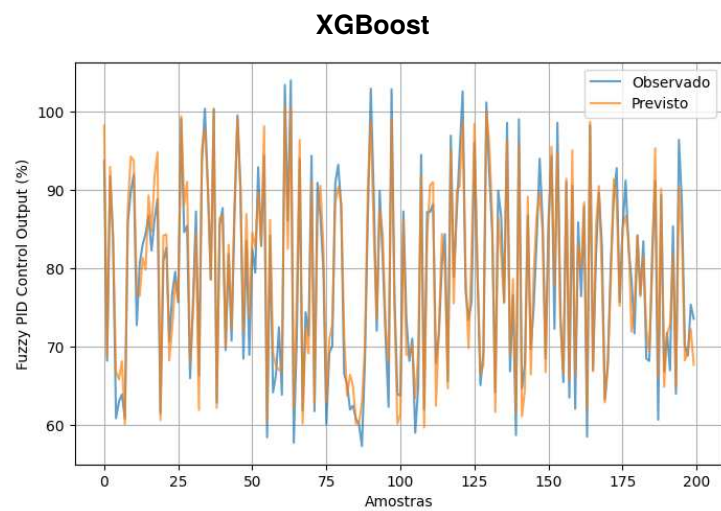


Figura 19 – Gráfico de saída real x prevista para o modelo XGBoost.

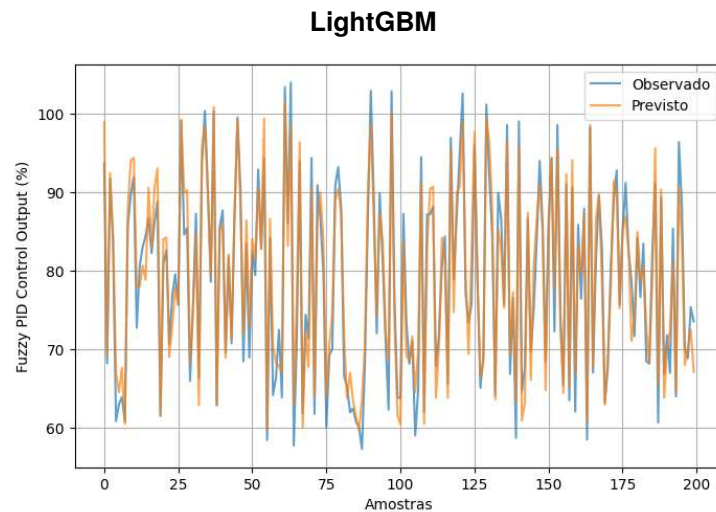


Figura 20 – Gráfico de saída real x prevista para o modelo LightGBM.

Em todos os modelos, os coeficientes de determinação (R^2) superiores a 0,8 resultam em boa aderência entre a curva de saída predita e a saída real, com grande sobreposição das linhas (geralmente representadas pelas cores laranja e azul). Essa sobreposição é mais evidente nos modelos lineares e de ensemble, que apresentam menor dispersão dos pontos ao redor da linha ideal.

No caso do SVR e do MLP, embora ainda exista correlação visível entre a saída real e a predita, observa-se uma maior discrepância em determinadas regiões, o que se manifesta pela maior exposição dos pontos correspondentes à saída real e pela menor coincidência entre as curvas. Esse comportamento é coerente com os valores mais altos de MSE, RMSE e MAE discutidos nas subseções anteriores e evidencia que esses modelos, na configuração adotada, não são os mais indicados para o problema proposto.

4.7 Síntese dos resultados e implicações práticas

De maneira geral, os resultados obtidos indicam que:

- As regressões Linear, Ridge e Lasso apresentam desempenho muito semelhante e superior aos demais modelos em todas as métricas analisadas (R^2 , MSE, RMSE e MAE);
- Modelos de ensemble baseados em árvores (Random Forest, Gradient Boosting, XGBoost e LightGBM) oferecem desempenho competitivo, porém sem superar as regressões lineares;

- SVR e MLP apresentaram os piores resultados relativos, com maiores erros e maior dispersão de resíduos.

Do ponto de vista prático, os modelos lineares destacam-se por aliarem bom desempenho preditivo e baixa complexidade computacional, além de maior interpretabilidade dos coeficientes. Em aplicações industriais, em que é desejável compreender a influência de cada variável de processo sobre a variável controlada, essa interpretabilidade é particularmente relevante.

Entre as regressões, a Lasso apresenta uma vantagem adicional ao possibilitar seleção de variáveis por meio da regularização L1, o que pode simplificar futuramente o modelo e reduzir custos de instrumentação, caso algumas variáveis de entrada se mostrem dispensáveis. Assim, os resultados obtidos apontam a Regressão Lasso como candidata natural a modelo de referência para a etapa de controle ou otimização do sistema estudado.

5 CONCLUSÕES

Esta pesquisa atingiu seus objetivos ao aplicar, comparar e validar modelos de Aprendizagem de Máquina para a previsão da ação de controle ideal em um sistema térmico baseado em Controlador Fuzzy-PID. Partindo da limitação clássica dos controladores PID, cuja eficácia depende fortemente da correta sintonia dos ganhos Proporcional (K_p), Integral (K_i) e Derivativo (K_d), o estudo demonstrou que técnicas modernas de regressão podem atuar como um mecanismo eficiente para apoiar e otimizar esse processo. Em sistemas com comportamento não linear, histerese térmica e variações ambientais — características comuns em processos industriais — a determinação de um conjunto ótimo de ganhos representa um desafio recorrente e de alto impacto operacional.

A introdução do controlador Fuzzy-PID já constitui um avanço ao agregar adaptabilidade e capacidade de ajuste dinâmico, compensando incertezas e irregularidades inerentes ao processo. No entanto, ao integrar Aprendizagem de Máquina ao problema da estimação da ação de controle ideal, este trabalho evidencia que modelos supervisionados podem não apenas replicar, mas também aprender o comportamento implícito da lógica fuzzy, capturando padrões estruturais que muitas vezes permanecem ocultos em análises tradicionais de controle.

Entre os nove modelos treinados, o algoritmo de Regressão Lasso apresentou o melhor desempenho global, atingindo um Coeficiente de Determinação ($R^2 = 0.9407$), o menor Erro Quadrático Médio (MSE) e os menores erros absolutos entre todas as abordagens testadas. Esse resultado revela que, embora o processo físico de controle térmico seja intrinsecamente não linear, a relação entre os parâmetros de entrada (erro, overshoot, tempo de resposta, temperatura ambiente, ganhos PID, entre outros) e a ação final aplicada pelo controlador Fuzzy-PID segue uma estrutura suficientemente regular para ser capturada por um modelo linear regularizado.

Esse achado possui implicações diretas e relevantes para a teoria e prática do controle:

- A relação entre variáveis de processo e ação de controle é mais previsível do que o esperado, permitindo que modelos lineares aprendam com alta precisão o comportamento de um controlador adaptativo.
- A regularização L1 do Lasso favorece a seleção automática de variáveis

relevantes, revelando quais parâmetros exercem maior influência sobre o controle — potencialmente simplificando futuras estratégias de modelagem ou instrumentação.

- Modelos de regressão podem auxiliar, acelerar ou até substituir métodos tradicionais de sintonia de controladores PID, reduzindo a dependência de heurísticas manuais e diminuindo o tempo de calibração.
- A integração de Aprendizagem de Máquina com estratégias de controle fortalece a adoção de soluções alinhadas aos princípios da Indústria 4.0, ampliando a automação inteligente e a autonomia operacional.

De forma geral, o estudo confirma o papel da Aprendizagem de Máquina como ferramenta de engenharia capaz de internalizar estratégias de controle adaptativas, oferecendo alternativas computacionalmente eficientes para o complexo problema da sintonia ótima de controladores. Os resultados demonstram que Machine Learning não apenas otimiza, mas pode também expandir as capacidades de sistemas clássicos, proporcionando previsibilidade, redução de custos e maior robustez em ambientes industriais.

Trabalhos Futuros

Com base nos resultados obtidos, recomenda-se a continuidade da pesquisa nas seguintes direções:

- Explorar técnicas mais avançadas de otimização de hiperparâmetros, como *Bayesian Optimization*, Optuna ou busca adaptativa.
- Investigar modelos híbridos que combinem Aprendizagem de Máquina com Controle Preditivo Baseado em Modelo (MPC).
- Expandir a base de dados com diferentes condições operacionais, perturbações externas e cenários reais de indústria.

Conclui-se que a abordagem proposta apresenta alto potencial de aplicação prática, contribuindo para a evolução de sistemas de controle inteligentes e para o avanço tecnológico alinhado às demandas da Indústria 4.0.

REFERÊNCIAS

AGUADO, Alexandre Garcia; CANTANHEDE, Marco André. Lógica fuzzy. **Artigo sem**, 2010.

ALCÂNTARA JUNIOR, Gilberto Pereira de. Avaliação do lasso e métodos alternativos em modelos de regressão logística, 2021.

ALEXANDRE, G et al. Comparativo entre controladores PID, PID Fuzzy e Preditivo avaliados a una planta industrial de nível e temperatura. In: CONGRESSO Técnico Científico da Engenharia e da Agronomia CONTECC', nº Paraná, Brasil. [S. l.: s. n.], 2016.

ALVES, Laurizio Emanuel Ribeiro; GOMES, Heliofábio Barros. Validação da imputação múltipla via Predictive Mean Matching para preenchimento de falhas nos dados pluviométricos da bacia do médio São Francisco. **Anuário do Instituto de Geociências**, v. 43, n. 1, p. 199–206, 2020.

ANTUNES, Ricardo. **Uberização, trabalho digital e indústria 4.0**. [S. l.]: Boitempo Editorial, 2020.

AWAD, Mariette; KHANNA, Rahul. Support vector regression. In: EFFICIENT learning machines: Theories, concepts, and applications for engineers and system designers. [S. l.]: Springer, 2015. p. 67–80.

BREIMAN, Leo. Random forests. **Machine learning**, Springer, v. 45, n. 1, p. 5–32, 2001.

CAI, Yuliang et al. Smart Nutrient Solution Temperature Control System for Oversummering Lettuce Cultivation Based on Adaptive Dung Beetle Optimizer-Fuzzy PID. **Applied Sciences**, v. 15, n. 10, p. 5381, 2025. DOI: 10.3390/app15105381. Disponível em: <https://www.mdpi.com/2076-3417/15/10/5381>.

CASTRO MELO, Carlen Rodrigues de. **ANÁLISE DO PROCESSO DE MONTAGEM DE PLACAS DE CIRCUITO IMPRESSO DE NOTEBOOK EM UMA EMPRESA DO POLO INDUSTRIAL DE MANAUS**. 2017. Tese (Doutorado) – Universidade Federal do Pará.

CHAI, Tianfeng; DRAXLER, Roland R et al. Root mean square error (RMSE) or mean absolute error (MAE). **Geoscientific model development discussions**, v. 7, n. 1, p. 1525–1534, 2014.

CHEN, Tianqi; GUESTRIN, Carlos. Xgboost: A scalable tree boosting system. In: PROCEEDINGS of the 22nd acm sigkdd international conference on knowledge discovery and data mining. [S. l.: s. n.], 2016. p. 785–794.

CHIODE, Angelo de Souza. **Avaliação de propostas de coeficientes de determinação do tipo R^2 em modelos de regressão logística com resposta nominal**. 2021. Tese (Doutorado) – Universidade de São Paulo.

DRAPER, NR. **Applied regression analysis**. [S. l.]: McGraw-Hill. Inc, 1998.

GAO, Yuan; ZHU, Wanshan. A Segmented Adaptive PID Temperature Control Method Suitable for Industrial Dispensing System. **Electronics**, MDPI, v. 14, n. 11, p. 2306, 2025.

HÖRBE, Roman; EROL, Selim. Optimizing LightGBM for Regression: A Study on Parameter Influence and Performance. **IFAC-PapersOnLine**, Elsevier, v. 59, n. 10, p. 2292–2297, 2025.

IZBICKI, Rafael; SANTOS, Tiago Mendonça dos. **Aprendizado de máquina: uma abordagem estatística**. [S. l.]: Rafael Izbicki, 2020.

JUNIOR, Marco Antonio Berni et al. Implementação de um Controlador Fuzzy para Controle de Temperatura. **Brazilian Journal of Development**, v. 6, n. 6, p. 38231–38245, 2020.

LI, Wenfang; WANG, Yuqiao. Intelligent Temperature Control Method of Instrument Based on Fuzzy PID Control Technology. **International Journal of Advanced Computer Science & Applications**, v. 15, n. 1, 2024.

MARRO, Alessandro Assi et al. Lógica fuzzy: conceitos e aplicações. **Natal: Universidade Federal do Rio Grande do Norte (UFRN)**, v. 2, 2010.

MATOS, Manuel António. Manual operacional para a regressão linear. **Faculdade de Engenharia da Universidade do Porto**, v. 63, 1995.

MEDINA, Mário Sérgio Martinelli et al. MODELAGEM PREDITIVA DE IMPRESSÕES COM SKLEARN E GRADIENT BOOSTING. **REVISTA DIÁLOGO E INTERAÇÃO**, v. 18, n. 1, p. 535–551, 2024.

MONARD, Maria Carolina; BARANAUSKAS, José Augusto. Conceitos sobre aprendizado de máquina. **Sistemas inteligentes-Fundamentos e aplicações**, v. 1, n. 1, p. 32, 2003.

MONTGOMERY, Douglas C; PECK, Elizabeth A; VINING, G Geoffrey. **Introduction to linear regression analysis**. [S. l.]: John Wiley & Sons, 2021.

OGATA, Katsuhiko. **Modern control engineering**. [S. l.]: Prentice hall, 2010.

POPESCU, Marius-Constantin et al. Multilayer perceptron and neural networks. **WSEAS Transactions on Circuits and Systems**, v. 8, n. 7, p. 579–588, 2009.

REYNALDO, Cristiane. **Regressão"e; Ridge"e;: um metodo alternativo para o mal condicionamento da matriz das regressoras**. 1997. Tese (Doutorado) – [sn].

SACOMANO, José Benedito et al. **Indústria 4.0**. [S. l.]: Editora Blucher, 2018.