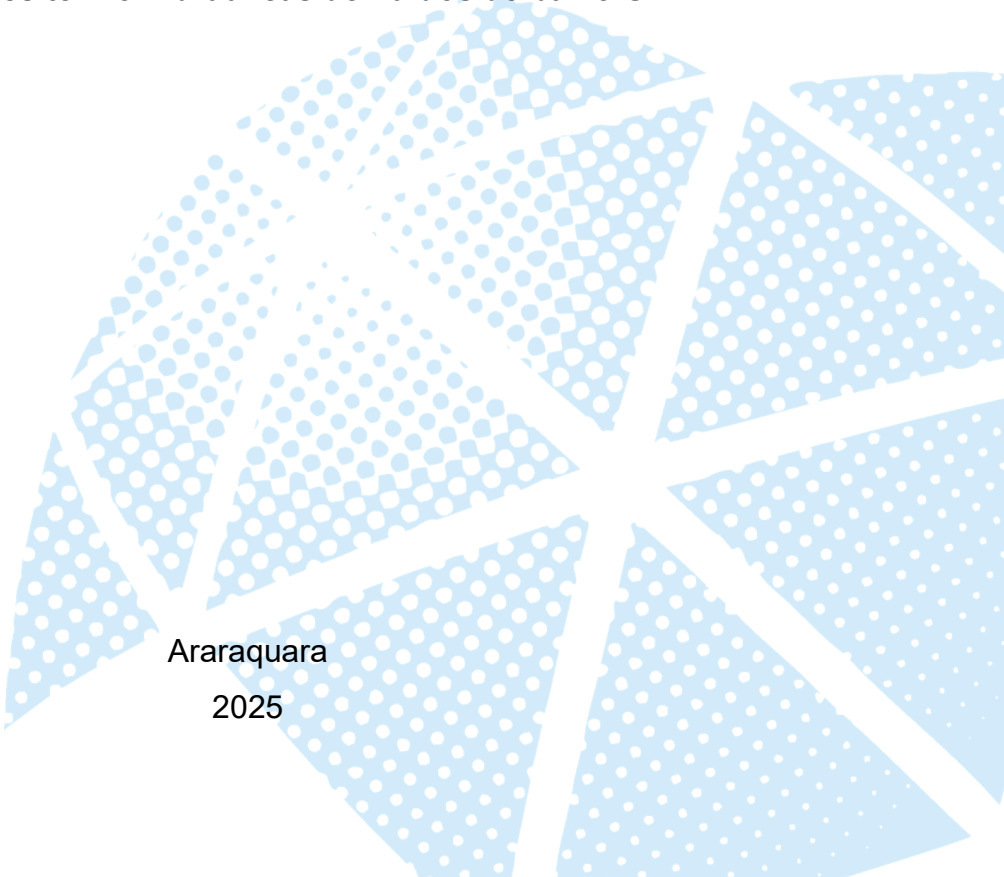


UNIVERSIDADE ESTADUAL PAULISTA – UNESP
Instituto de Química - Câmpus de Araraquara

FÁBIO CARVALHO SCAREL

**Estudo comparativo entre modelos de aprendizado de máquina para predição
de propriedades termo-hidráulicas de fluidos de baixo GWP.**

Araraquara
2025



FÁBIO CARVALHO SCAREL

Estudo comparativo entre modelos de aprendizado de máquina para predição de propriedades termo-hidráulicas de fluidos de baixo GWP.

Trabalho de Conclusão de Curso
apresentado à Universidade Estadual
Paulista (UNESP), Instituto de Química,
Araraquara, para obtenção do título de
Bacharel em Engenharia Química.

Orientador(a): Profa. Dra. Érica Regina
Filletti Nascimento

Araraquara

2025

S285e Scarel, Fábio

Estudo comparativo entre modelos de aprendizado de máquina para predição de propriedades termo-hidráulicas de fluidos de baixo GWP. / Fábio Scarel. -- Araraquara, 2025
41 p.

Trabalho de conclusão de curso (Bacharelado - Engenharia Química) - Universidade Estadual Paulista (UNESP), Instituto de Química, Araraquara
Orientadora: Érica Regina Filleti

1. Machine learning. 2. Fluidos refrigerantes. 3. Engenharia química. 4. Proteção ambiental. I. Título.

FÁBIO CARVALHO SCAREL

Estudo comparativo entre modelos de aprendizado de máquina para predição de propriedades termo-hidráulicas de fluidos de baixo GWP.

Trabalho de Conclusão de Curso
apresentado à Universidade Estadual
Paulista (UNESP), Instituto de Química,
Araraquara, para obtenção do título de
Bacharel em Engenharia Química.

Data da defesa: 08/12/2025

Banca Examinadora:

Profa. Dra. Érica Regina Filletti Nascimento

Prof. Dra. Marisa Veiga Capela

Prof. Dr. Adriano do Santos

AGRADECIMENTOS

Gostaria de agradecer aos meus pais Márcia Garcia, José Scarel e meu irmão Pedro Scarel pelo suporte durante todos os momentos. Gostaria de agradecer minha namorada Laura Dias por me apoiar e acompanhar nas aventuras da vida.

Agradecer também aos meus amigos Paulo Beltrami, Rodrigo Tarallo, João Borin, Vitor Sartorello, Ian Guedes Godinho, João Pedro Ramos e Thiago Baptistella por promoverem todas as discussões enriquecedoras que tivemos.

Agradeço também a Profa. Dra. Érica Nascimento pela oportunidade de estudar programação e algoritmos de machine learning dedicados a aplicações na Engenharia Química.

Finalmente ao Instituto de Química de Araraquara por todos os recursos disponibilizados e aprendizados que obtive durante meus anos como aluno.

RESUMO

Este trabalho investiga a aplicação de técnicas de aprendizado de máquina para prever propriedades termo-hidráulicas de fluidos refrigerantes de baixo Potencial de Aquecimento Global (GWP) durante a ebulição convectiva em microcanais. Utilizaram-se os dados experimentais de Sempértegui-Tapia (2016), compostos pelos fluidos R134a, R1234ze(E), R1234yf e R600a em geometrias circular, quadrada e triangular, incluindo medições de perda de pressão por atrito e coeficiente de troca térmica em diferentes condições operacionais. O conjunto de dados foi pré-processado com interpolação de propriedades termofísicas e cálculo de números adimensionais, seguidos de padronização e codificação binária das variáveis categóricas. Foram avaliados três algoritmos representativos de abordagens amplamente utilizadas em dados tabulares: redes neurais Multilayer Perceptron (MLP), SAINT (Self-Attention and Intersample Transformer) e o modelo baseado em árvores de decisão XGBoost. A comparação de desempenho foi realizada com métricas quantitativas como erro médio absoluto (EMA), erro médio relativo (MRE) e porcentagem de acerto dentro de margens de $\pm 20\%$ e $\pm 30\%$. Os resultados demonstraram que o XGBoost obteve o melhor desempenho geral para ambas as propriedades analisadas, reduzindo o erro significativamente em relação às correlações empíricas reportadas na literatura experimental. O MLP apresentou desempenho satisfatório e relativamente próximo ao XGBoost em certas condições, enquanto o modelo SAINT não convergiu adequadamente para o volume de dados disponível, resultando em erros elevados e baixa acurácia. Os achados evidenciam que modelos de ensemble baseados em árvores constituem alternativas robustas e precisas para a previsão do comportamento termo-hidráulico em microcanais, contribuindo para o desenvolvimento de sistemas de refrigeração mais eficientes e ambientalmente sustentáveis. Conclui-se que todos os objetivos definidos foram plenamente atendidos, com a identificação clara do algoritmo mais eficaz para a modelagem estudada.

Palavras-chave: fluidos de baixo GWP; ebulição convectiva; microcanais; aprendizado de máquina; XGBoost.

ABSTRACT

This work investigates the application of machine learning techniques to predict thermohydraulic properties of low Global Warming Potential (GWP) refrigerants during convective boiling in microchannels. Experimental data provided by Sempértegui-Tapia (2016) were used, including the refrigerants R134a, R1234ze(E), R1234yf, and R600a in circular, square, and triangular channel cross-sections, containing measurements of frictional pressure drop and heat transfer coefficient under various operating conditions. The dataset was pre-processed through interpolation of thermophysical properties, calculation of dimensionless numbers, feature standardization, and one-hot encoding of categorical variables. Three representative algorithms for tabular data were evaluated: a Multilayer Perceptron (MLP) neural network, the SAINT (Self-Attention and Intersample Transformer) architecture, and the XGBoost gradient-boosted decision-tree model. Quantitative performance comparison relied on error metrics such as Mean Absolute Error (MAE), Mean Relative Error (MRE), and the percentage of predictions within $\pm 20\%$ and $\pm 30\%$ error bands. The results demonstrated that XGBoost achieved the best overall performance for both target variables, significantly reducing prediction error relative to empirical correlations available in the experimental literature. The MLP showed satisfactory performance, at times close to XGBoost, while SAINT failed to converge adequately due to the limited dataset size, resulting in large prediction errors. The findings confirm that ensemble-based models constitute robust and accurate alternatives for predicting thermohydraulic behavior in microchannels, supporting the development of more efficient and environmentally sustainable refrigeration systems. All objectives established in this work were fully achieved, including the identification of the most suitable predictive algorithm for the studied phenomena.

Keywords: low-GWP refrigerants; convective boiling; microchannels; machine learning; XGBoost.

LISTA DE FIGURAS

Figura 1 - Evolução dos padrões de escoamento durante a ebulição convectiva em microcanais.	11
Figura 2 - Evolução dos padrões de escoamento durante a ebulição convectiva em canais de diâmetros regulares.	11
Figura 3 – Gráfico de violino das variáveis presentes no conjunto de dados completo de perda de pressão na tese original	14
Figura 4 – Gráfico de violino das variáveis presentes no conjunto de dados completo do coeficiente de troca térmica na tese original	14
Figura 5 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria circular de perda de pressão de na tese original	14
Figura 6 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria triangular de perda de pressão de na tese original.....	15
Figura 7 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria quadrada de perda de pressão de na tese original.....	15
Figura 8 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria circular do coeficiente de troca térmica na tese original	15
Figura 9 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria triangular do coeficiente de troca térmica na tese original.....	15
Figura 10 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria quadrada do coeficiente de troca térmica na tese original.....	16
Figura 11 - Separação em treino/validação/teste para o conjunto de dados de perda de pressão.....	21
Figura 12 - Separação em treino/validação/teste para o conjunto de dados de perda de pressão.....	21
Figura 13 – Representação de um Multilayer Perceptron.....	23
Figura 14 – Exemplo de funcionamento do SAINT para o conjunto MNIST	26
Figura 15 - Distribuição do erro relativo por modelo	30
Figura 16 – Comparação entre os dados esperados e os dados previstos em escala logarítmica.....	30
Figura 17 – Distribuição dos erros de acordo com cada modelo.....	31
Figura 18 - Comparação entre os dados esperados e os previstos em escala logarítmica.....	31
Figura 19 - Erro médio absoluto por modelo, geometria e fluido	35

Figura 20 - Erro médio absoluto por modelo, geometria e fluido	36
Figura 21 - Comparação entre os valores medidos e os valores previstos pelo XGBoost para o coeficiente de troca térmica	37
Figura 22 - Comparação entre os valores medidos e os valores previstos pelo XGBoost para a perda de pressão	38

LISTA DE TABELAS

Tabela 1 – Número de amostras por fluido e geometria para perda de pressão por atrito	17
Tabela 2 – Propriedades da Tabela 1 e suas unidades	17
Tabela 3 – Número de amostras por fluido e geometria para coeficiente de troca de calor	17
Tabela 4 – Propriedades da Tabela 3 e suas unidades	17
Tabela 5 - XGBoost – Perda de pressão	32
Tabela 6 - MLP - Perda de Pressão.....	33
Tabela 7 - SAINT - Perda de Pressão	33
Tabela 8 – XGBoost – Coeficiente de troca térmica	34
Tabela 9 – MLP – Coeficiente de troca térmica	34
Tabela 10 - SAINT - Coeficiente de troca térmica.....	34

SUMÁRIO

1 INTRODUÇÃO	10
2 OBJETIVOS	12
2.1 OBJETIVOS ESPECÍFICOS	12
3 MATERIAIS E MÉTODOS	13
3.1 CONJUNTO DE DADOS.....	13
3.2 PRÉ-PROCESSAMENTO	16
3.2.1 Extração	16
3.2.2 Interpolação e tratamento das propriedades termofísicas	17
3.2.3 Números adimensionais	18
3.2.4 Separação em conjunto de treinamento, validação e teste	20
3.3 ARQUITETURAS DE MACHINE LEARNING AVALIADAS	22
3.3.1 MLP (Multi-Layer Perceptron).....	22
3.3.2 SAINT	25
3.3.3 XGBoost	27
4 RESULTADOS.....	28
4.1 ANÁLISE EXPLORATÓRIA DOS RESÍDUOS E ESTATÍSTICAS DESCRITIVAS	29
4.1.1 Perda de pressão	29
4.1.2 Coeficiente de troca térmica	30
4.2 DESEMPENHO GERAL DOS MODELOS.....	31
4.2.1 Perda de Pressão	32
4.2.2 Coeficiente de troca térmica	33
4.3 DESEMPENHO POR FLUIDO E GEOMETRIA	34
4.3.1 Perda de pressão	35
4.3.2 Coeficiente de troca térmica	35
4.4 DISCUSSÃO COMPARATIVA.....	36
5 CONCLUSÃO.....	38

1 INTRODUÇÃO

O desenvolvimento de tecnologias energeticamente eficientes e ambientalmente sustentáveis tem se intensificado nas últimas décadas diante dos desafios impostos pelas mudanças climáticas globais. No setor de refrigeração, essa transição é particularmente relevante, visto que os fluidos refrigerantes tradicionais apresentam elevado Potencial de Aquecimento Global (GWP), contribuindo significativamente para o efeito estufa (IPCC, 2021).

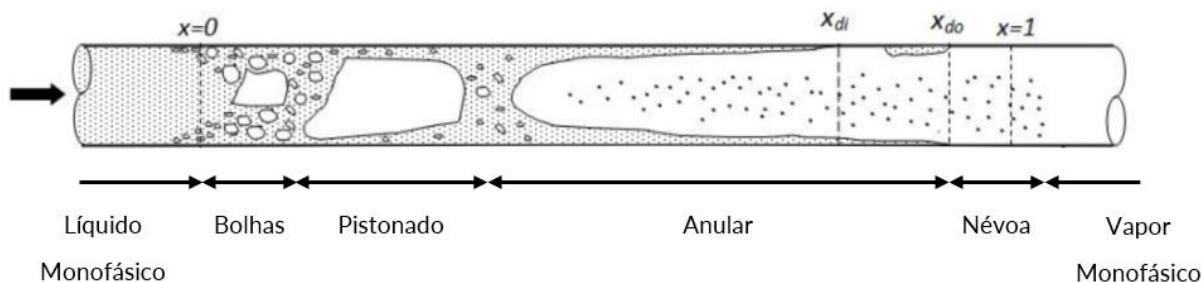
Nesse contexto, diversos países e organizações internacionais vêm promovendo a substituição de substâncias de alto impacto ambiental por alternativas de baixo GWP, como os hidrofluoroolefinas (HFOs), R1234yf e R1234ze(E) e os hidrocarbonetos naturais, como o R600a (isobutano) (BOLAJI; HESKETH; ADEKUNLE, 2014; LOMBARDO et al., 2025). Tais fluidos representam alternativas promissoras na busca por sistemas de refrigeração mais sustentáveis, conciliando eficiência energética e redução das emissões diretas de gases de efeito estufa.

Apesar das vantagens ambientais, a aplicação desses novos fluidos em trocadores de calor compactos caracterizados por canais de dimensões reduzidas impõe desafios técnicos significativos. A ebulição convectiva em microcanais apresenta comportamentos altamente complexos, influenciados por fatores como propriedades termofísicas do fluido, regime de escoamento, fluxo de calor e geometria da seção transversal. A variação geométrica (circular, quadrada e triangular) e as condições operacionais (pressão de saturação, vazão mássica, título e temperatura) exercem influência direta sobre o coeficiente de transferência de calor e a perda de carga por atrito, dificultando a generalização de correlações empíricas tradicionais.

A ebulição convectiva é um processo de transferência de calor em sistemas bifásicos no qual um fluido em escoamento forçado passa por uma superfície aquecida, promovendo simultaneamente o movimento do fluido e a vaporização. Nesse regime, o transporte de calor resulta da combinação entre a convecção impulsionada pelo escoamento e a mudança de fase líquido-vapor, sendo influenciado por variáveis como calor aplicado, fluxo de massa e propriedades termofísicas do fluido. No contexto de microcanais, que são canais com diâmetro hidráulico tipicamente inferior a 1 mm, as características de ebulição convectiva diferem das observadas em canais maiores, devido a efeitos de confinamento geométrico, maior razão superfície/volume e diferentes regimes de fluxo bifásico que influenciam a

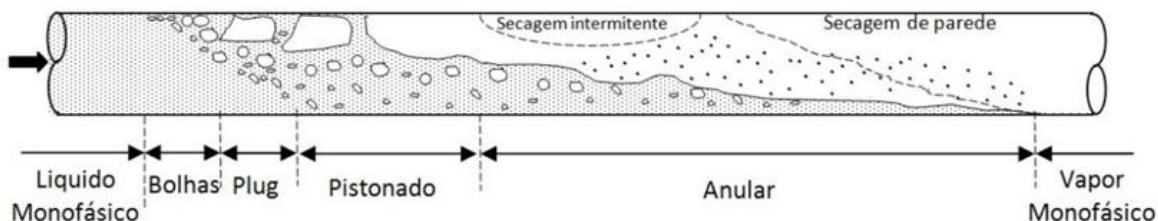
formação de bolhas e a dinâmica do filme líquido ao longo da parede aquecida. Essas diferenças exigem correlações e modelos específicos para microcanais, pois os métodos desenvolvidos para canais convencionais nem sempre capturam adequadamente esses comportamentos em micro-escala.

Figura 1 - Evolução dos padrões de escoamento durante a ebulição convectiva em microcanais.



Fonte: Adaptado de Sempertégui-Tapia (2016)

Figura 2 - Evolução dos padrões de escoamento durante a ebulição convectiva em canais de diâmetros regulares.



Fonte: Adaptado de Collier e Thome (1996)

Estudos experimentais, como o de Sempertégui-Tapia (2016), forneceram importantes contribuições ao mapeamento do comportamento térmico e hidráulico desses fluidos em diferentes configurações geométricas. No entanto, as correlações empíricas desenvolvidas a partir de tais dados tendem a apresentar limitações de extrapolação quando aplicadas a novos fluidos, geometrias ou faixas operacionais, reduzindo sua capacidade preditiva.

Nesse cenário, técnicas de aprendizado de máquina (machine learning) emergem como alternativas promissoras, uma vez que são capazes de capturar relações não lineares complexas e identificar padrões ocultos em grandes conjuntos de dados experimentais. Modelos como as redes neurais Multilayer Perceptron (MLP), arquiteturas baseadas em atenção como o SAINT (Self-Attention and Intersample Transformer) e métodos de ensemble como o XGBoost (Extreme Gradient Boosting)

têm demonstrado desempenho competitivo em problemas tabulares. (SOMEPELLI et al., 2021; CHEN; GUESTRIN, 2016; ZOU et al., 2023).

Dessa forma, este trabalho tem como objetivo desenvolver e avaliar modelos de aprendizado de máquina capazes de prever o coeficiente de troca térmica e a perda de carga por atrito durante a ebulição convectiva em canais de dimensões reduzidas. Para isso, são utilizados os dados experimentais obtidos por Sempértegui-Tapia (2016) como base de treinamento e validação, considerando os fluidos R134a, R1234ze(E), R1234yf e R600a em diferentes geometrias de seção transversal (circular, quadrada e triangular).

Ao comparar o desempenho de arquiteturas baseadas em redes neurais e modelos de árvores de decisão, busca-se identificar a metodologia de maior acurácia e robustez, contribuindo para o aprimoramento de modelos preditivos aplicáveis ao dimensionamento de trocadores de calor mais eficientes e sustentáveis.

2 OBJETIVOS

O principal objetivo deste trabalho foi desenvolver modelos de redes neurais artificiais MLP, SAINT e um modelo de árvores XGBoost capazes de estimar o coeficiente de troca térmica e a perda de pressão por atrito dos fluidos R134a, R1234ze(E), R1234yf e R600a em canais de dimensões reduzidas.

2.1 Objetivos específicos

Os objetivos específicos desse trabalho foram:

- Implementar modelos baseados em aprendizado de máquina (MLP, SAINT e XGBoost) para previsão do coeficiente de troca térmica e da perda de pressão por atrito de fluidos refrigerantes de baixo GWP em microcanais;
- Comparar o desempenho das diferentes arquiteturas quanto à precisão, robustez e capacidade de generalização, identificando o modelo mais adequado para representar o comportamento térmico e hidráulico dos fluidos estudados;
- Definir o melhor algoritmo de aprendizagem visando a menor porcentagem de erro para o conjunto de dados;

3 MATERIAIS E MÉTODOS

Nesta seção será explicado a origem dos dados e o tratamento que foi feito neles para ser utilizado pelos algoritmos de aprendizagem, além de uma breve explicação de cada um dos métodos testados

3.1 Conjunto de dados

Esta seção descreve as etapas metodológicas empregadas neste estudo, incluindo a origem e caracterização dos dados experimentais, o pré-processamento realizado, e a implementação dos modelos de aprendizado de máquina utilizados para a previsão do coeficiente de troca térmica e da perda de carga por atrito durante a ebulição convectiva em canais de dimensões reduzidas.

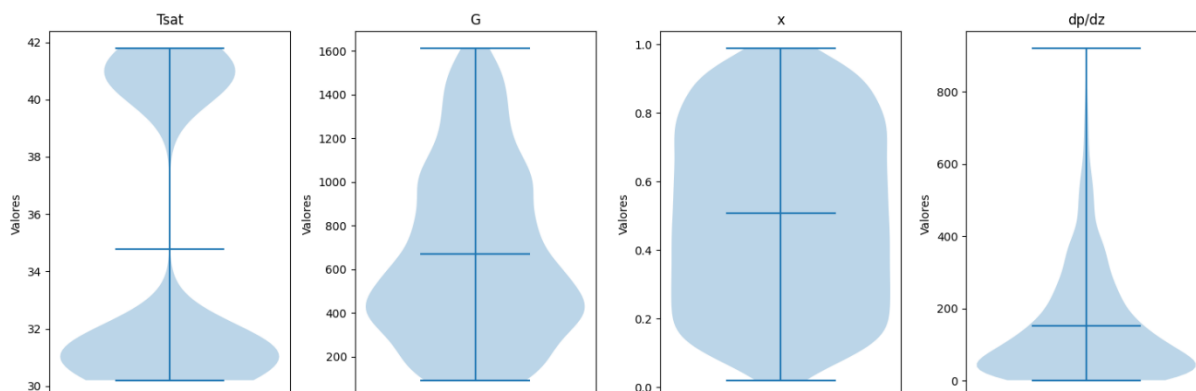
O conjunto de dados utilizado neste trabalho foi obtido a partir da tese de doutorado de Sempértegui-Tapia (2016), defendida na Escola de Engenharia de São Carlos (EESC-USP). O banco de dados contém resultados experimentais de escoamento bifásico com ebulição convectiva em microcanais, abrangendo quatro fluidos refrigerantes de baixo GWP: R134a, R1234ze(E), R1234yf e R600a.

As medições foram realizadas para três geometrias distintas de seção transversal (circular, quadrada e triangular) totalizando 1468 observações de perda de carga por atrito e 4579 observações de coeficiente de troca de calor.

Esses dados incluem variáveis experimentais como temperatura de saturação (T_{sat}), fluxo mássico (G), título de vapor (x), fluxo de calor aplicado (ϕ) e coeficiente de transferência de calor (h), além da perda de pressão por atrito (dp/dz)_f.

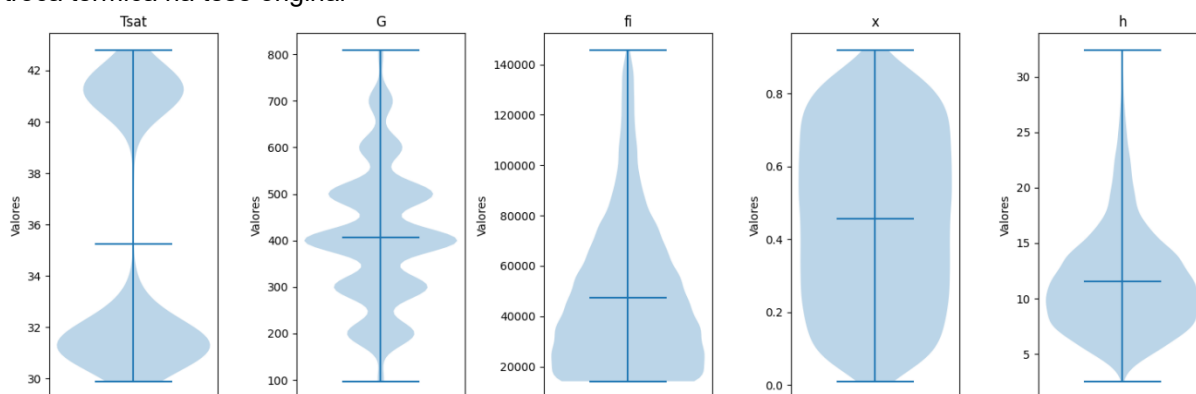
As Figuras 1 a 8 ilustram a distribuição estatística das variáveis envolvidas no conjunto experimental, evidenciando a não uniformidade das medições e a predominância de determinadas faixas operacionais. Essa análise preliminar reforça a necessidade do uso de modelos capazes de capturar relações não lineares complexas, pois o comportamento termo-hidráulico é altamente dependente da interação entre variáveis.

Figura 3 – Gráfico de violino das variáveis presentes no conjunto de dados completo de perda de pressão na tese original



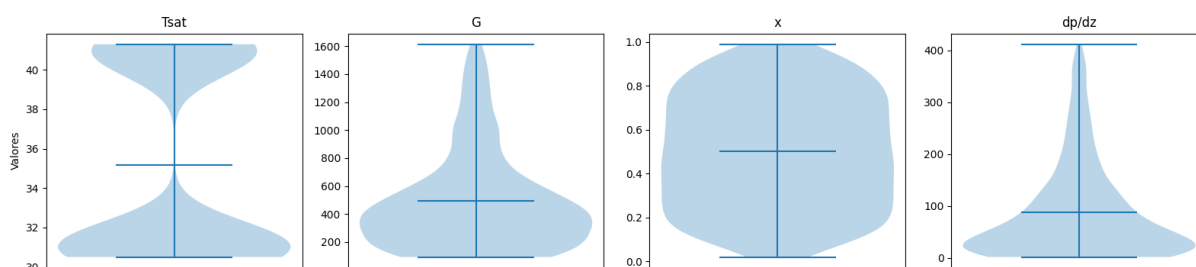
Fonte: Elaborado pelo próprio autor

Figura 4 – Gráfico de violino das variáveis presentes no conjunto de dados completo do coeficiente de troca térmica na tese original



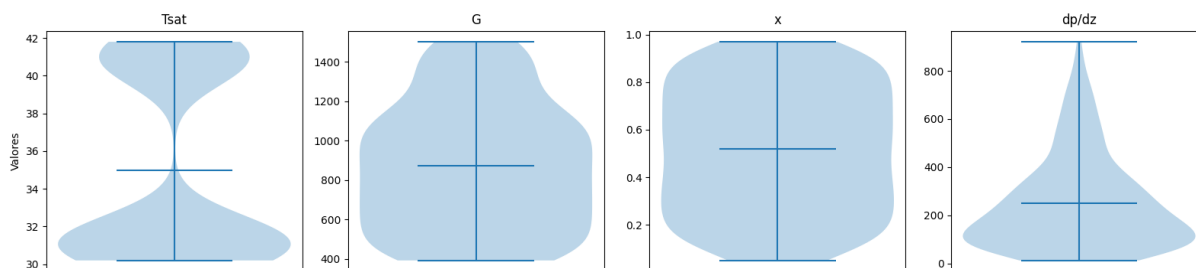
Fonte: Elaborado pelo próprio autor

Figura 5 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria circular de perda de pressão de na tese original



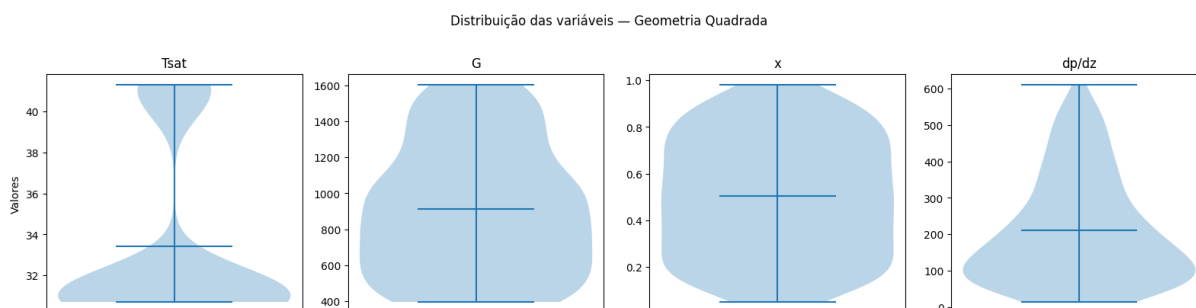
Fonte: Elaborado pelo próprio autor

Figura 6 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria triangular de perda de pressão de na tese original



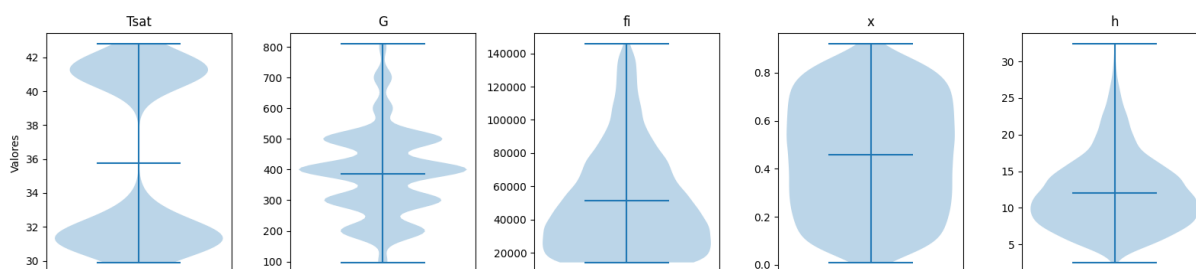
Fonte: Elaborado pelo próprio autor

Figura 7 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria quadrada de perda de pressão de na tese original



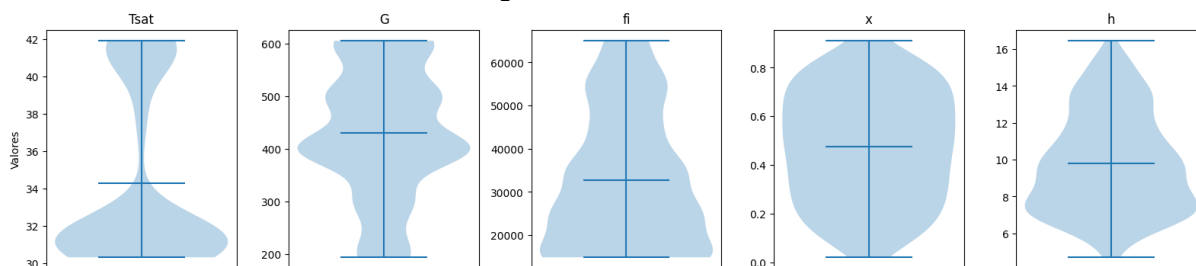
Fonte: Elaborado pelo próprio autor

Figura 8 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria circular do coeficiente de troca térmica na tese original



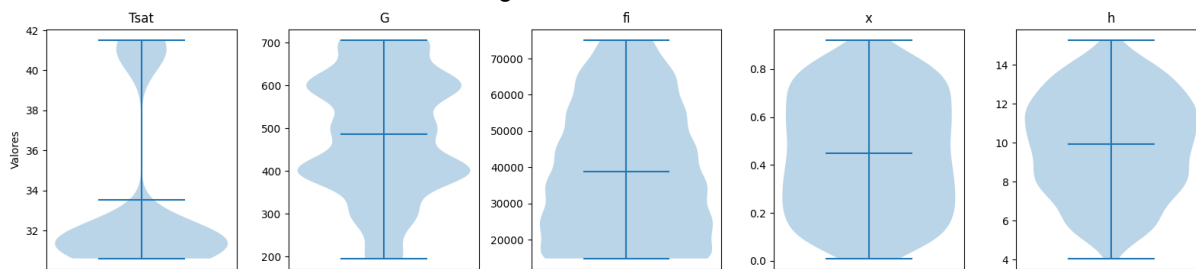
Fonte: Elaborado pelo próprio autor

Figura 9 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria triangular do coeficiente de troca térmica na tese original



Fonte: Elaborado pelo próprio autor

Figura 10 - Gráfico de violino das variáveis presentes no conjunto de dados com geometria quadrada do coeficiente de troca térmica na tese original



Fonte: Elaborado pelo próprio autor

3.2 Pré-processamento

O pré-processamento dos dados constitui uma etapa fundamental para garantir que as variáveis experimentais apresentem consistência física, compatibilidade numérica e qualidade estatística adequadas para aplicação dos modelos de aprendizado supervisionado.

Algumas propriedades fundamentais, como densidades, viscosidades dinâmicas e tensão superficial, não estavam diretamente disponíveis na tese original portanto foram buscadas em outras literaturas (BOLAJI et al; LOMBARDO et al; ZHAO et al; LEMMON et al; BELL et al). Assim, métodos de interpolação e regressões polinomiais foram empregados para completar e tornar contínua a base de propriedades termofísicas, possibilitando o cálculo dos números adimensionais essenciais na representação do escoamento bifásico.

As variáveis numéricas do conjunto de dados foram submetidas a um processo de normalização por meio do método de padronização, no qual cada variável é transformada de modo a apresentar média nula e variância unitária. Esse procedimento tem como objetivo reduzir os efeitos de diferenças de escala entre os atributos, evitando que variáveis com maior ordem de grandeza exerçam influência desproporcional no processo de treinamento dos modelos.

3.2.1 Extração

As tabelas 1 a 4 descrevem os dados originais extraídos para da tese. Para a perda de pressão:

Tabela 1 – Número de amostras por fluido e geometria para perda de pressão por atrito

Contagem	Fluido	Geometria
253	R134a	Circular
186	R600a	Circular
229	R1234ze(E)	Circular
140	R1234yf	Circular
295	R134a	Quadrada
365	R134a	Triangular

Fonte: Elaborado pelo próprio autor

Tabela 2 – Propriedades da Tabela 1 e suas unidades

Propriedade	Unidade
Tsat	°C
G	kg/m ² s
x	-
(dp/dz) _f	kPa/m

Fonte: Elaborado pelo próprio autor

Para o coeficiente de troca de calor:

Tabela 3 – Número de amostras por fluido e geometria para coeficiente de troca de calor

Contagem	Fluido	Geometria
772	R134a	Circular
1315	R600a	Circular
847	R1234ze(E)	Circular
474	R1234yf	Circular
753	R134a	Quadrada
418	R134a	Triangular

Fonte: Elaborado pelo próprio autor

Tabela 4 – Propriedades da Tabela 3 e suas unidades

Propriedade	Unidade
Tsat	°C
G	kg/m ² s
φ	kW/m ²
x	-
h	kW/m ² K

Fonte: Elaborado pelo próprio autor

3.2.2 Interpolação e tratamento das propriedades termofísicas

Sempértegui-Tapia (2016) utiliza de números adimensionais para estudar o escoamento bifásico, que serão valiosos para os algoritmos de aprendizado.

Para calculá-los foram utilizadas outras propriedades dos fluidos estudados, como a densidade e a viscosidade dinâmica do líquido e do gás, assim como a tensão superficial.

As propriedades básicas (densidade, viscosidade dinâmica e tensão superficial) foram obtidas a partir do banco de dados NIST Chemistry WebBook (LEMMON et al., 2025) e complementadas com dados da literatura (ZHAO et al., 2014). O tratamento foi realizado em Python 3.10, utilizando as bibliotecas Pandas, NumPy, Scikit-learn e CoolProp (BELL et al., 2014).

Os dados fornecidos por ZHAO et al. estavam em faixas específicas que não contemplavam todo o conjunto de dados de Sempértegui-Tapia (2016), portando foi realizada uma regressão polinomial de grau 2 para obter os valores em todas as faixas de densidade e viscosidade do líquido e do gás.

Em seguida, aplicou-se uma interpolação polinomial de segunda ordem, obtida pelo método dos mínimos quadrados, para expressar a variação da densidade (ρ) e viscosidade (μ) em função da temperatura de saturação (T_{sat}). Essa regressão foi implementada por meio da função `numpy.polyfit`.

A tensão superficial (σ) foi determinada a partir da função `PropsSI` da biblioteca CoolProp, considerando a condição de saturação do líquido ($Q = 0$) e a temperatura de operação de cada medição.

Esse procedimento assegurou a homogeneidade e completude do conjunto de dados de entrada, permitindo que as variações das propriedades termofísicas fossem tratadas de forma contínua e coerente ao longo das faixas de temperatura, título e geometria. Como resultado, a base final apresentou consistência física e numérica adequada para a aplicação dos modelos de aprendizado de máquina.

Com as propriedades termofísicas devidamente interpoladas, foram calculados os números adimensionais relevantes ao escoamento bifásico.

3.2.3 Números adimensionais

Número de Reynolds

O número de Reynolds (Re) caracteriza a razão entre as forças inerciais e viscosas do escoamento. Para escoamentos bifásicos, ele é definido separadamente para as fases líquida e gasosa, conforme as Equações (1) e (2):

$$Re_L = \frac{G * D * (1 - x)}{\mu_L} \quad (1)$$

$$Re_G = \frac{G * D * x}{\mu_G} \quad (2)$$

Onde G é o fluxo mássico, D é o diâmetro do canal, x é a fração de vapor e μ_G é a viscosidade dinâmica do fluido em fase gasosa e μ_L é a viscosidade dinâmica do fluido em fase líquida.

Número de Weber

O número de Weber (We) expressa a razão entre as forças de inércia e as forças de tensão superficial, sendo dado por:

$$We_L = \frac{G^2 * D * (1 - x)}{\rho_L * \sigma} \quad (3)$$

$$We_G = \frac{G^2 * D * x}{\rho_G * \sigma} \quad (4)$$

Onde G é o fluxo mássico, D é o diâmetro do canal, x é a fração de vapor e ρ_G é a densidade do fluido em fase gasosa e ρ_L é a densidade do fluido em fase líquida e σ é a tensão superficial.

Esse número é amplamente utilizado na caracterização de escoamentos bifásicos e na previsão da formação de ondas e instabilidades interfaciais (CANDAL et al., 2017).

Número de confinamento

O grau de confinamento (Co) das bolhas com relação as dimensões de um canal é medido por esse número adimensional e ele é dado por:

$$Co = \frac{[\sigma / ((\rho_L - \rho_G)g)]^{1/2}}{L} \quad (5)$$

Onde g é a aceleração da gravidade e L é o comprimento do espaço confinador, que no caso de canais com seção transversal circular ou um polígono regular, é dado pelo diâmetro hidráulico.

Número de Suratman

O número de Suratman (Su) combina os efeitos viscosos e de tensão superficial, sendo uma função dos números de Reynolds e de Weber:

$$Su_L = \frac{Re_L^2}{We_L} \quad (6)$$

$$Su_G = \frac{Re_G^2}{We_G} \quad (7)$$

O número de Suratman pode ser definido para as duas fases e relaciona efeitos de tensão superficial e viscosos. Para fases líquidas, números de Suratman reduzidos indicam formação de ondas.

Número de Froude

O número de Froude (Fr) representa a relação entre forças de inércia e gravitacionais:

$$Fr_L = \frac{G^2 * (1 - x)}{\rho_G^2 * g * D} \quad (8)$$

Em escoamentos líquidos - gás horizontais o número de Froude relaciona-se a efeitos de estratificação, os quais podem ser relevantes inclusive para canais de diâmetro reduzido.

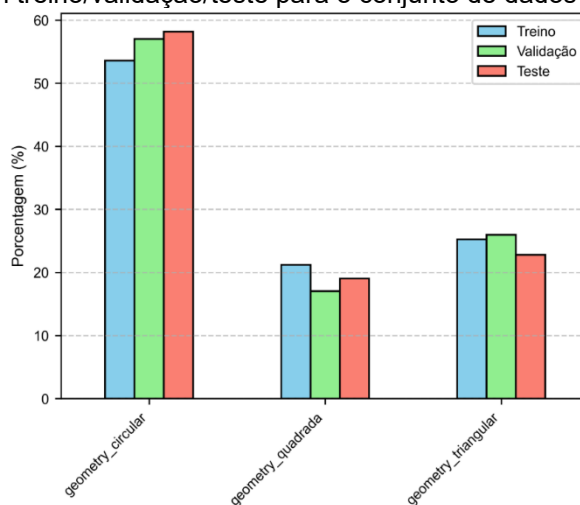
3.2.4 Separação em conjunto de treinamento, validação e teste

Com o objetivo de avaliar de forma rigorosa a capacidade de generalização dos modelos de aprendizado supervisionado e evitar o fenômeno de *overfitting*, o conjunto experimental foi particionado em três subconjuntos independentes: treinamento, validação e teste. Essa abordagem segue as boas práticas recomendadas em ciência de dados e modelagem estatística, garantindo que o desempenho reportado represente a habilidade dos modelos em prever valores não utilizados durante o ajuste de parâmetros.

A divisão foi realizada empregando a função `train_test_split` da biblioteca `scikit-learn`, com embaralhamento aleatório das observações e estabelecimento de uma semente fixa (*random seed*) para assegurar a reprodutibilidade dos resultados. Na

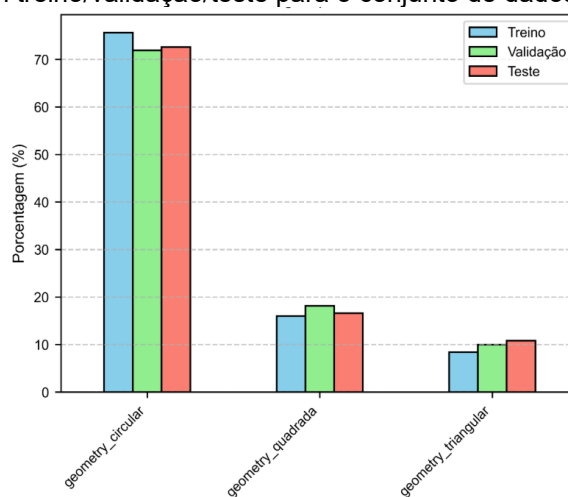
primeira etapa, 80% das observações totais foram alocadas ao conjunto de treinamento, enquanto 20% restantes foram destinados ao conjunto de teste. Em seguida, o conjunto de treinamento passou por uma segunda divisão, na qual 80% foram mantidos para treinamento e 20% destinados à validação do modelo ao longo do processo de ajuste de hiperparâmetros. A separação dos dados de acordo com a geometria pode ser vista pelas Figuras 9 e 10.

Figura 11 - Separação em treino/validação/teste para o conjunto de dados de perda de pressão



Fonte: elaborado pelo próprio autor

Figura 12 - Separação em treino/validação/teste para o conjunto de dados de perda de pressão



Fonte: elaborado pelo próprio autor

Em seguida, variáveis numéricas foram padronizadas por meio do StandardScaler da biblioteca scikit-learn, transformando-as para média zero e desvio-padrão unitário, o que favorece a estabilidade do treinamento em modelos sensíveis à escala, como redes neurais. Para as variáveis categóricas, restritas à identificação

da geometria da seção transversal do canal, utilizou-se One-Hot Encoding da biblioteca scikit-learn, convertendo cada categoria (circular, quadrada e triangular) em atributos binários independentes e garantindo que essas categorias fossem distribuídas igualmente e pudessem ser validadas em todas as etapas sem distorções.

Além disso, o ajuste de parâmetros da normalização feita por StandardScaler foi realizado apenas no conjunto de treinamento e aplicado em validação e teste, evitando contaminação de dados externos ao treinamento.

Finalmente, os dados referentes aos fluidos foram excluídos para não interferir no treinamento dos modelos.

3.3 Arquiteturas de Machine Learning Avaliadas

Este capítulo descreve as três arquiteturas utilizadas no experimento: MLP, SAINT e XGBoost. Para cada modelo, são apresentados seus princípios de funcionamento, vantagens, limitações e aspectos relevantes para aplicação em dados tabulares.

As arquiteturas de aprendizado de máquina avaliadas neste trabalho foram implementadas em linguagem Python (versão 3.10). O modelo Multi-Layer Perceptron (MLP) foi desenvolvido utilizando o framework TensorFlow, o modelo XGBoost foi implementado por meio da biblioteca xgboost, e o modelo SAINT (Self-Attention and Intersample Attention Transformer) com a biblioteca lit-saint, disponível publicamente em ambiente open-source. Todas as execuções foram realizadas em ambiente Windows 11, com gerenciamento de dependências via ambiente virtual (.venv) e suporte a aceleração por GPU quando aplicável.

3.3.1 MLP (Multi-Layer Perceptron)

O modelo de rede neural artificial é inspirado no funcionamento do sistema nervoso biológico, no qual neurônios interconectados processam sinais elétricos e produzem respostas adaptativas. A primeira formulação matemática desse conceito foi apresentada por McCulloch e Pitts (1943), que descreveram o neurônio como uma unidade lógica binária capaz de representar funções booleanas, estabelecendo as bases teóricas para os sistemas conexionistas.

Décadas depois, Frank Rosenblatt (1958) introduziu o Perceptron, considerado o primeiro modelo de aprendizado supervisionado capaz de ajustar pesos sinápticos de forma automática. Seu trabalho, desenvolvido no Cornell Aeronautical Laboratory,

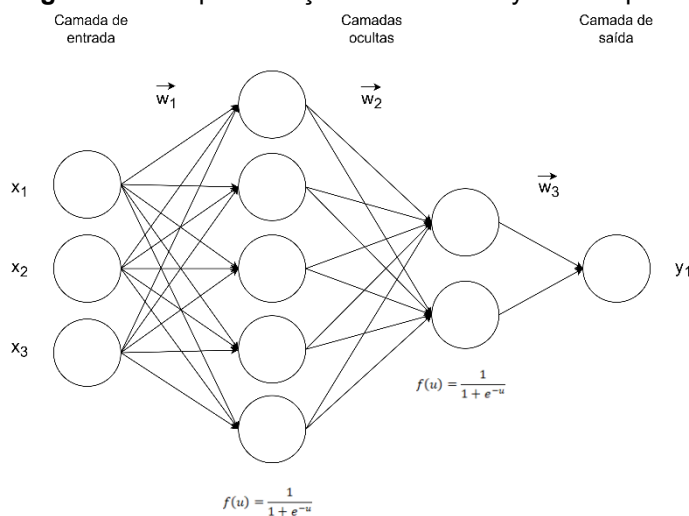
demonstrou que sistemas artificiais poderiam reconhecer padrões a partir de dados experimentais, ainda que restritos a uma única camada de processamento.

A superação dessas limitações ocorreu com o surgimento das redes multicamadas, conhecidas como Multi-Layer Perceptron (MLP). A proposta de redes compostas por múltiplas camadas de neurônios interconectados foi inicialmente discutida por Shun'ichi Amari (1967), que sugeriu o uso de métodos de otimização baseados em gradiente para o treinamento de redes profundas. Contudo, foi somente com o trabalho de Rumelhart, Hinton e Williams (1986) que o MLP se tornou viável na prática, graças à introdução do algoritmo de retropropagação do erro (backpropagation), o qual permitiu o ajuste eficiente dos pesos sinápticos em todas as camadas da rede.

O MLP pertence à categoria de redes feedforward, nas quais a informação flui em uma única direção, sem realimentação, como representado na Figura 11. A arquitetura típica é composta por três tipos de camadas:

- Camada de entrada: recebe as variáveis independentes (neste trabalho, parâmetros experimentais como fluxo mássico, título de vapor e temperatura de saturação);
- Camadas ocultas: responsáveis por processar combinações não lineares das entradas;
- Camada de saída: fornece a predição final

Figura 13 – Representação de um Multilayer Perceptron



Fonte: elaborado pelo próprio autor

Cada camada é formada por unidades de processamento chamadas neurônios artificiais, inspirados no funcionamento biológico dos neurônios humanos. Esses neurônios artificiais recebem valores de entrada x_i , associados a pesos sinápticos w_i , os quais representam a importância de cada entrada. O produto dessas entradas ponderadas é somado a um termo de bias (b), que permite deslocar a fronteira de decisão do modelo, já que sem ele o hiperplano definido pelo neurônio seria forçado a passar pela origem, conforme a Equação (9):

$$u = \sum_{i=1}^n w_i x_i + b \quad (9)$$

O valor resultante u é então processado por uma função de ativação $f(u)$, que introduz a não linearidade necessária para que a rede seja capaz de modelar relações complexas entre as variáveis de entrada e saída. As funções de ativação mais utilizadas são a sigmoide, a tangente hiperbólica e a ReLU (Rectified Linear Unit). A sigmoide é expressa pela Equação (10):

$$f(u) = \frac{1}{1 + e^{-u}} \quad (10)$$

Essas funções controlam a propagação dos sinais pela rede e ajudam a estabilizar o aprendizado, limitando a faixa de saída dos neurônios.

O processo de aprendizado de uma MLP ocorre de maneira supervisionada, isto é, a rede é treinada com exemplos cujas saídas reais (rótulos) são conhecidas. O objetivo do treinamento é ajustar os pesos sinápticos de modo que a saída produzida pela rede (y_{pred}) se aproxime o máximo possível da saída real (y_{real}).

Para isso, utiliza-se um algoritmo de otimização baseado no gradiente descendente. Esse algoritmo calcula o erro quadrático médio (MSE) entre as saídas previstas e as reais, conforme a Equação (12):

$$E = \frac{1}{N} \sum_{j=1}^N (y_{\text{real},j} - y_{\text{pred},j})^2 \quad (11)$$

Em seguida, o erro é propagado de forma inversa pela rede, atualizando os pesos de cada camada de acordo com o gradiente do erro em relação aos pesos. Essa atualização é feita conforme a Equação (12):

$$w_{ij}^{(t+1)} = w_{ij}^{(t)} - \eta \frac{\partial E}{\partial w_{ij}} \quad (12)$$

Em aplicações práticas, as MLPs são amplamente utilizadas em processos industriais, previsão de séries temporais, análise de dados experimentais e modelagem de propriedades físico-químicas, sendo especialmente eficazes quando as relações entre as variáveis não seguem padrões lineares evidentes (SOUZA et al., 2012).

3.3.2 SAINT

O modelo SAINT (Self-Attention and Intersample Transformer) foi proposto por Somepalli et al. (2021) como uma nova abordagem de aprendizado profundo voltada para dados tabulares, utilizando mecanismos de atenção inspirados nos Transformers originalmente empregados em tarefas de linguagem natural. O principal objetivo dos autores foi adaptar a arquitetura Transformer, amplamente bem-sucedida em dados sequenciais, para cenários tabulares onde as amostras não possuem ordem temporal e as variáveis (colunas) exibem dependências complexas e heterogêneas.

Inicialmente, cada variável de entrada é representada por um vetor de *embedding*. Para atributos categóricos, utiliza-se uma camada de *embedding treinável*, enquanto os atributos numéricos passam por uma projeção linear para o mesmo espaço dimensional d_{model} . Assim, cada linha da base de dados, composta por D atributos, é convertida em uma sequência de tokens.

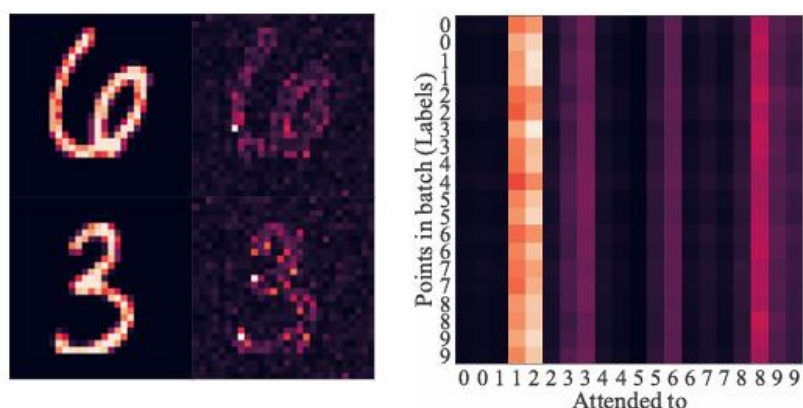
Diferentemente do Transformer convencional, o SAINT aplica dois tipos distintos de blocos de atenção, organizados de forma alternada ao longo da rede:

1. Atenção entre atributos (Intra-sample attention): Nesse estágio, o modelo aplica o mecanismo de self-attention sobre as colunas de uma única amostra, permitindo que cada atributo aprenda a se relacionar com os demais atributos da mesma linha. Essa etapa captura interdependências internas da amostra, identificando, por exemplo, como uma variável pode influenciar outra.
2. Atenção entre amostras (Inter-sample attention): Aqui, o modelo aplica atenção cruzada entre diferentes instâncias do conjunto de dados, de forma que amostras semelhantes possam compartilhar informações relevantes. Essa

etapa melhora a generalização, especialmente em bases com alta variabilidade entre exemplos.

Esses dois módulos de atenção são aplicados alternadamente, o que permite ao SAINT construir representações enriquecidas que levam em conta tanto as relações horizontais (entre atributos) quanto as relações verticais (entre amostras) da base de dados.

Figura 14 – Exemplo de funcionamento do SAINT para o conjunto MNIST



Fonte: SOMEPALI et al., 2021

(a) na esquerda representa a relação entre colunas da mesma amostra no algoritmo SAINT e (b) representa como o modelo enxerga as relações entre amostras diferentes.

As Figuras 12(a) e 12(b) exemplificam o funcionamento dos mecanismos de atenção utilizados no SAINT.

Na Figura 12(a), observa-se a autoatenção (self-attention) aplicada a amostras do conjunto MNIST, em que o modelo aprende a identificar as regiões mais relevantes de cada imagem para o reconhecimento do dígito. As áreas mais claras representam maior peso de atenção, indicando os pixels que mais contribuíram para a decisão do modelo.

Já a Figura 12(b) ilustra o mecanismo de atenção entre amostras (intersample attention), no qual o modelo analisa relações entre diferentes instâncias dentro de um mesmo lote (batch). Cada linha representa uma amostra e cada coluna indica o grau de atenção direcionado às demais observações. As regiões mais intensas revelam que o SAINT tende a concentrar a atenção em exemplos com rótulos semelhantes, demonstrando sua capacidade de capturar similaridades contextuais entre instâncias durante o processo de aprendizado (SOMEPALI et al., 2021).

3.3.3 XGBoost

O XGBoost (Extreme Gradient Boosting) é um algoritmo de aprendizado supervisionado proposto por Chen e Guestrin (2016), amplamente reconhecido por sua alta eficiência e desempenho em problemas de classificação e regressão com dados tabulares.

Ele pertence à família dos métodos de ensemble boosting, que combinam múltiplos modelos fracos (geralmente árvores de decisão) para formar um modelo robusto e preciso.

Diferentemente dos métodos de bagging (como Random Forests), que treinam várias árvores de forma independente e agregam seus resultados por média ou votação, o boosting segue uma abordagem sequencial e aditiva, em que cada nova árvore é treinada para corrigir os erros das árvores anteriores.

O princípio do *gradient boosting* consiste em minimizar uma função de perda diferenciável por meio de uma soma aditiva de modelos. Dado um conjunto de dados com n amostras (x_i, y_i) , o modelo é construído de forma incremental:

$$\hat{y}_i^{(t)} = \sum_{k=1}^t f_k(x_i), \quad f_k \in \mathcal{F} \quad (13)$$

onde cada f_k representa uma árvore de decisão e $\mathcal{F}$ é o espaço de todas as árvores possíveis. A cada iteração t , adiciona-se uma nova árvore que aproxima o gradiente negativo da função de perda em relação às previsões atuais, ajustando o modelo de forma semelhante ao gradiente descendente.

A função objetivo geral é dada por:

$$\mathcal{L}^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (14)$$

em que:

- $l(y_i, \hat{y}_i)$ é a função de perda (por exemplo, erro quadrático médio ou entropia cruzada);
- $\Omega(f_t)$ é um termo de regularização que controla a complexidade da nova árvore.

Expandindo a função de perda e ignorando termos constantes, obtém-se:

$$\mathcal{L}^{(t)} = \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \quad (15)$$

onde:

- $g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$ representa o gradiente de primeira ordem;
- $h_i = \partial_{\hat{y}_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)})$ representa o gradiente de segunda ordem

Um dos grandes diferenciais do XGBoost é a inclusão explícita de regularização L1 e L2 no termo $\Omega(f_t)$, o que reduz o *overfitting* e melhora a generalização. Esse termo é definido como:

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (16)$$

onde:

- T é o número de folhas da árvore;
- w_j é o peso da folha j ;
- γ penaliza a adição de novas folhas (complexidade estrutural);
- λ controla a penalização L2 sobre os pesos (complexidade paramétrica).

Durante o crescimento das árvores, o XGBoost realiza a divisão (split) apenas quando a redução esperada na função de perda supera o custo imposto pela regularização — evitando, assim, a criação de divisões que não agregam ganho informativo relevante.

4 RESULTADOS

Esta seção apresenta e discute os resultados obtidos a partir da aplicação dos modelos de aprendizado de máquina avaliados neste estudo para a previsão do coeficiente de troca térmica e da perda de pressão durante a ebulição convectiva em microcanais. Inicialmente, realiza-se uma análise exploratória dos resíduos, observando o comportamento estatístico das diferenças entre os valores medidos e previstos pelos modelos. Em seguida, é apresentada uma avaliação quantitativa do desempenho global das arquiteturas implementadas, baseada em métricas padronizadas de erro e acurácia.

Posteriormente, os resultados são segmentados por fluido refrigerante e geometria de seção transversal, permitindo investigar a robustez dos modelos sob diferentes condições operacionais. Por fim, os comportamentos observados são analisados de maneira comparativa, destacando as vantagens e limitações de cada abordagem no contexto termo-hidráulico estudado.

4.1 Análise exploratória dos resíduos e estatísticas descritivas

A avaliação dos resíduos apresentada nas Figuras 13 a 17 permite observar a forma como os modelos se comportam frente aos dados experimentais disponíveis após o treinamento. O foco desta análise consiste em interpretar diretamente os padrões visualizados nos gráficos, descrevendo as características observáveis das distribuições de erro e da comparação entre valores previstos e reais, sem inferências sobre comportamento físico ou faixas operacionais além daquelas explicitamente representadas nas figuras.

4.1.1 Perda de pressão

Na Figura 13, os histogramas apresentam diferenças claras no comportamento do erro relativo entre os três modelos. O XGBoost demonstra maior concentração de valores próximos de zero, formando uma distribuição visualmente estreita e com pequena dispersão em torno do centro. Esse padrão sugere que, no conjunto de teste apresentado, suas previsões mantiveram uma proximidade mais consistente em relação aos valores medidos.

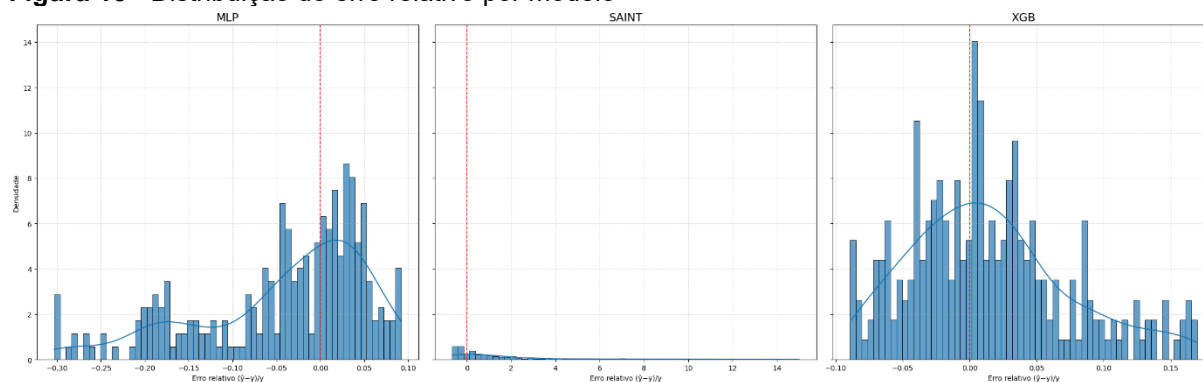
O MLP, por sua vez, exibe uma distribuição mais alargada, com erros espalhados em uma faixa mais ampla e com maior quantidade de barras distantes do centro, com um rabo para os erros relativos negativos, indicando que o modelo tende a subestimar os valores de perda de pressão.

O SAINT se diferencia dos demais: a Figura 13 mostra barras projetadas em regiões de erros extremamente elevados, com valores que ultrapassam significativamente aqueles apresentados pelo XGBoost e pelo MLP. A intensidade visual dessas barras torna evidente que o modelo apresentou resultados com erros muito maiores e frequentes, não conseguindo ser acertivo em suas previsões.

A Figura 14 reforça essas observações ao comparar previsões e valores experimentais em gráfico de dispersão. O XGBoost apresenta um agrupamento de pontos mais próximos da linha de tendência, com um R^2 de 0,98 comparado ao R^2 de

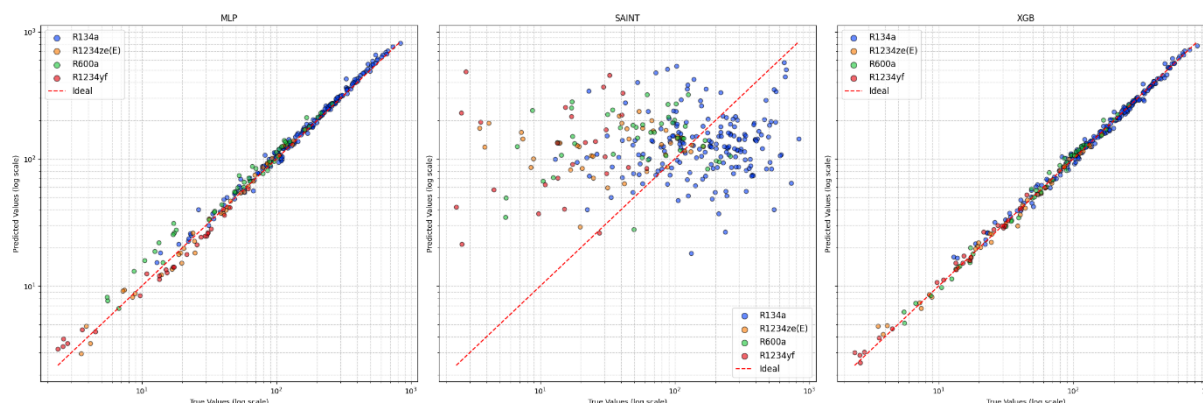
0,90 do MLP, sugerindo maior aderência na previsão. No MLP, os pontos se encontram visivelmente mais espalhados, enquanto o SAINT exibe uma dispersão extremamente acentuada e afastada da linha, o que confirma o desvio sistemático percebido anteriormente no histograma.

Figura 15 - Distribuição do erro relativo por modelo



Fonte: elaborado pelo próprio autor

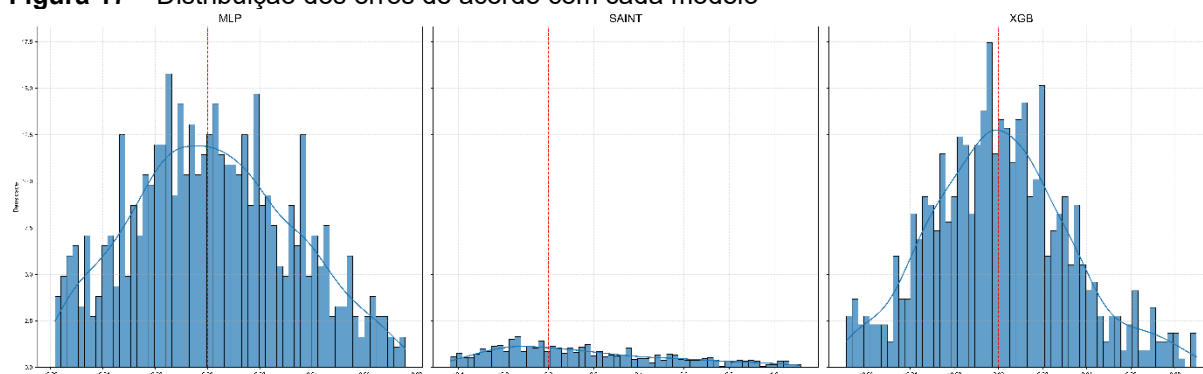
Figura 16 – Comparação entre os dados esperados e os dados previstos em escala logarítmica



Fonte: elaborado pelo próprio autor

4.1.2 Coeficiente de troca térmica

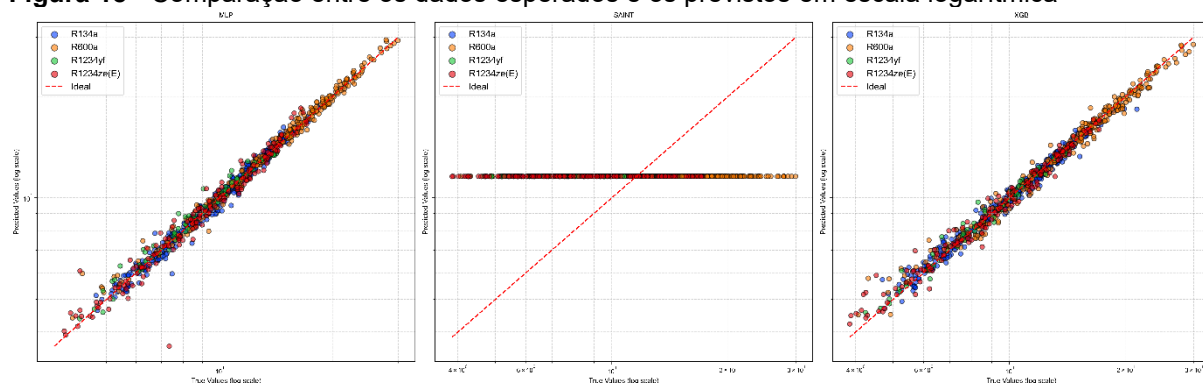
A Figura 15 indica distribuições de erro concentradas ao redor do zero para o MLP e para o XGBoost, com amplitude visualmente menor se comparada aos gráficos de perda de pressão. Essa concentração sugere que os valores previstos mantiveram maior proximidade em relação às medições experimentais. Com relação ao SAINT observamos novamente esse padrão de erros distribuídos por uma ampla faixa de valores. Isso indica que o modelo convergiu diminuindo o erro de todos os pontos e mas não conseguiu encontrar uma solução sólida.

Figura 17 – Distribuição dos erros de acordo com cada modelo

Fonte: elaborado pelo próprio autor

No gráfico de dispersão da Figura 16, os pontos associados ao XGBoost se alinham de maneira bastante próxima à linha de referência, enquanto o MLP também mostra pontos próximos a essa linha, ainda que com leve dispersão adicional quando comparado ao XGBoost. O R^2 confirma essa hipótese com valores próximos de 0,95 para ambos modelos.

Com relação ao SAINT, reforça-se a teoria que o modelo tentou convergir diminuindo de maneira igualitária o erro relativo de todos os pontos.

Figura 18 - Comparação entre os dados esperados e os previstos em escala logarítmica

Fonte: elaborado pelo próprio autor

4.2 Desempenho geral dos modelos

A análise quantitativa do desempenho dos modelos foi conduzida a partir dos indicadores apresentados nas Tabelas 5 a 10, que reportam os valores de erro médio absoluto (EMA), erro médio relativo (MRE) e as porcentagens de acertos dentro das margens de $\pm 30\%$ e $\pm 20\%$ (η_{30} e η_{20}) para os conjuntos de teste das duas propriedades avaliadas. Além disso, os resultados foram comparados diretamente

com aqueles obtidos na tese de Sempértégui-Tapia (2016), que utilizou os mesmos dados experimentais como base para o desenvolvimento de seu método preditivo.

4.2.1 Perda de Pressão

Observa-se que o XGBoost foi o modelo com melhor desempenho no presente estudo, apresentando EMA total de 5,9%, com 99,3% e 96,9% das previsões dentro dos limites de $\pm 30\%$ e $\pm 20\%$, respectivamente (Tabela 5). Esses resultados superam claramente os apresentados por Sempértégui-Tapia (2016), cujo método alcançou EMA total de 10,2%, com 97,2% e 89,2% das previsões dentro das mesmas faixas de tolerância. Assim, com base apenas nos valores apresentados nos quadros comparativos, verifica-se uma redução de aproximadamente 42% no erro absoluto médio e um aumento consistente nas previsões bem-sucedidas dentro das margens estabelecidas, indicando uma melhora significativa no desempenho preditivo para esta variável.

O MLP, por sua vez, apresentou EMA total de 9,5% (Tabela 6), com menores percentuais de acurácia dentro das faixas de tolerância quando comparado ao XGBoost, comportamento coerente com a maior dispersão dos resíduos observada na Figura 13. Ainda assim, seu desempenho se mantém semelhante ao método de Sempértégui-Tapia em alguns fluidos e geometrias, embora inferior no conjunto global.

O SAINT apresentou desempenho bastante inferior para a perda de pressão, com EMA total de 349,8% (Tabela 7), resultado que confirma a instabilidade observada graficamente nas dispersões da Figura 12 e evidencia que o modelo não convergiu de forma adequada para os dados utilizados.

Tabela 5 - XGBoost – Perda de pressão

Base de dados	Geometria	# Dados	EMA (%)	MRE (%)	η_{30} (%)	η_{20} (%)
R1234yf	circular	26	7,3	5,0	100,0	92,3
R1234ze(E)	circular	39	6,1	0,6	97,4	94,9
R134a	circular	61	6,7	2,5	98,4	95,1
R134a	quadrada	56	5,2	2,2	100,0	98,2
R134a	triangular	67	4,4	-0,3	100,0	100,0
R600a	circular	45	7,2	2,7	100,0	97,8
Total		294	5,9	1,8	99,3	96,9

Fonte: elaborado pelo próprio autor

Tabela 6 - MLP - Perda de Pressão

Base de dados	Geometria	# Dados	EMA (%)	MRE (%)	η_{30} (%)	η_{20} (%)
R1234yf	circular	26	8,7	-0,4	96,2	92,3
R1234ze(E)	circular	39	8,8	-5,5	92,3	87,2
R134a	circular	61	7,7	4,6	98,4	96,7
R134a	quadrada	56	8,1	7,8	98,2	96,4
R134a	triangular	67	11,9	11,8	95,5	89,6
R600a	circular	45	11,4	11,2	91,1	80,0
Total		294	9,5	6,1	95,6	90,8

Fonte: elaborado pelo próprio autor

Tabela 7 - SAINT - Perda de Pressão

Base de dados	Geometria	# Dados	EMA (%)	MRE (%)	η_{30} (%)	η_{20} (%)
R1234yf	circular	26	1679,0	1679,0	3,8	3,8
R1234ze(E)	circular	39	617,9	614,2	17,9	10,3
R134a	circular	61	108,2	66,8	34,4	21,3
R134a	quadrada	56	111,1	47,5	16,1	8,9
R134a	triangular	67	68,1	-18,6	13,4	4,5
R600a	circular	45	393,6	378,3	13,3	8,9
Total		294	349,8	306,5	18,0	10,2

Fonte: elaborado pelo próprio autor

4.2.2 Coeficiente de troca térmica

Em relação ao coeficiente de troca térmica, a Tabela 8 mostra novamente a superioridade do XGBoost, alcançando EMA total de 3,4%, com 99,8% e 99,2% das previsões dentro das margens de $\pm 30\%$ e $\pm 20\%$ (η_{30} e η_{20}), respectivamente. Esses resultados superam de forma expressiva aqueles apresentados por Sempértegui-Tapia, que reportou EMA total de 11,3%, com 97,5% das previsões dentro de $\pm 30\%$ e 85,3% dentro de $\pm 20\%$. Assim, no caso do coeficiente de transferência de calor, a redução relativa do erro em comparação ao estudo original ultrapassa 70%, representando um avanço substancial na precisão preditiva da variável.

O MLP apresentou desempenho consistente, com EMA total de 3,3% (Tabela 9), em linha com o XGBoost e superior ao método de Sempértegui-Tapia em todos os indicadores de tolerância, o que confirma o padrão já observado na Figura 14, em que os dados previstos se mantêm próximos da reta de referência. O SAINT, embora menos discrepante que no caso da perda de pressão, ainda apresentou resultados inferiores aos demais modelos, com maior erro médio e menor percentual de acerto dentro das faixas especificadas (Tabela 10).

Tabela 8 – XGBoost – Coeficiente de troca térmica

Base de dados	Geometria	# Dados	EMA (%)	MRE (%)	η_{30} (%)	η_{20} (%)
R1234yf	circular	86	2,7	1,0	100,0	100,0
R1234ze(E)	circular	167	4,2	1,0	99,4	97,6
R134a	circular	147	2,4	-0,1	100,0	100,0
R134a	quadrada	152	2,5	0,5	100,0	100,0
R134a	triangular	99	4,1	0,6	100,0	99,0
R600a	circular	265	3,9	-0,2	99,6	99,2
Total		916	3,4	0,3	99,8	99,2

Fonte: elaborado pelo próprio autor

Tabela 9 – MLP – Coeficiente de troca térmica

Base de dados	Geometria	# Dados	EMA (%)	MRE (%)	η_{30} (%)	η_{20} (%)
R1234yf	circular	86	3,3	1,8	100,0	100,0
R1234ze(E)	circular	167	4,6	0,2	98,8	98,8
R134a	circular	147	2,4	0,3	100,0	100,0
R134a	quadrada	152	2,7	1,3	100,0	100,0
R134a	triangular	99	3,6	-1,9	100,0	99,0
R600a	circular	265	3,1	1,2	99,6	99,6
Total		916	3,3	0,6	99,7	99,6

Fonte: elaborado pelo próprio autor

Tabela 10 - SAINT - Coeficiente de troca térmica

Base de dados	Geometria	# Dados	EMA (%)	MRE (%)	η_{30} (%)	η_{20} (%)
R1234yf	circular	86	42,3	38,4	53,5	46,5
R1234ze(E)	circular	167	48,5	39,0	50,3	37,1
R134a	circular	147	31,3	14,3	64,6	42,9
R134a	quadrada	152	33,1	26,4	62,5	55,9
R134a	triangular	99	37,8	31,7	56,6	41,4
R600a	circular	265	36,0	-8,0	43,8	23,4
Total		916	37,8	18,5	53,7	38,5

Fonte: elaborado pelo próprio autor

4.3 Desempenho por fluido e geometria

A análise segmentada por fluido e geometria permite observar com maior clareza variações de desempenho entre condições experimentais distintas. As Figuras 17 e 18 apresentam os valores de erro médio absoluto (EMA) obtidos para cada combinação analisada, permitindo comparar visualmente o desempenho relativo dos modelos dentro de cada subgrupo da base de dados. Essa observação direta

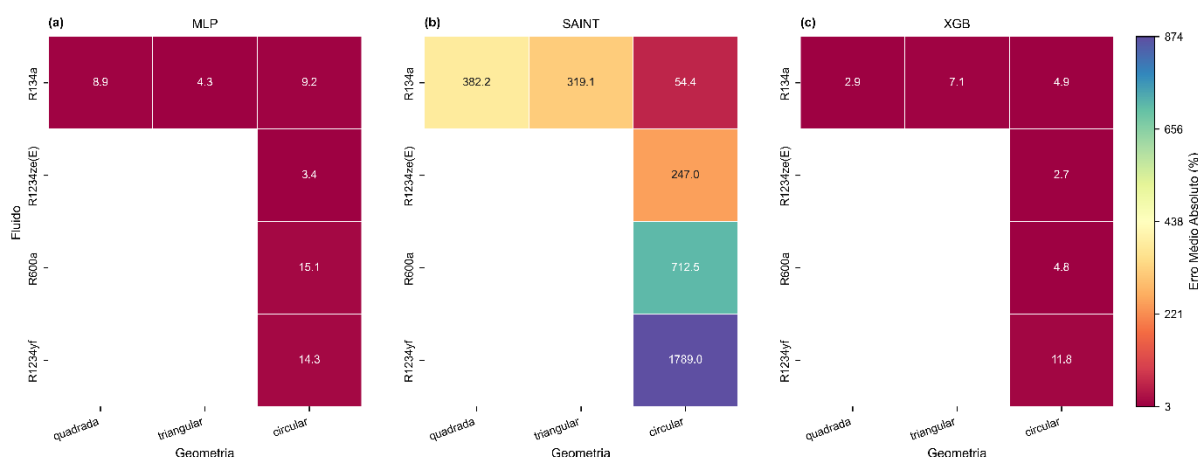
complementa a análise global reportada anteriormente, destacando áreas específicas em que os algoritmos demonstraram maior ou menor precisão.

4.3.1 Perda de pressão

Na Figura 17, observa-se que o XGBoost mantém sistematicamente os menores valores de erro médio absoluto para todos os fluidos e geometrias apresentados. O modelo apresenta desempenho consistente tanto para R134a (circular, quadrada e triangular) quanto para os fluidos R600a, R1234ze(E) e R1234yf na geometria circular. Uma performance parecida do MLP, entretanto com valores menores em todos os grupos, exceto o fluido R134a na geometria circular, onde ele teve uma acurácia maior.

Com relação ao SAINT e a geometria circular enxerga-se uma clara relação entre a quantidade de pontos dos fluidos e sua acurácia. Para o conjunto de dados completo (Tabela 1), o R134yf teve a menor quantidade de pontos (140,0) e o maior erro (1789,0%), enquanto o R134a teve a maior quantidade de dados (253,0) e o menor erro (54,0%). Isso pode ser explicado por sua arquitetura, se tratando de um modelo de deep learning, quanto maior a quantidade de pontos, melhor são ajustados os pesos.

Figura 19 - Erro médio absoluto por modelo, geometria e fluido



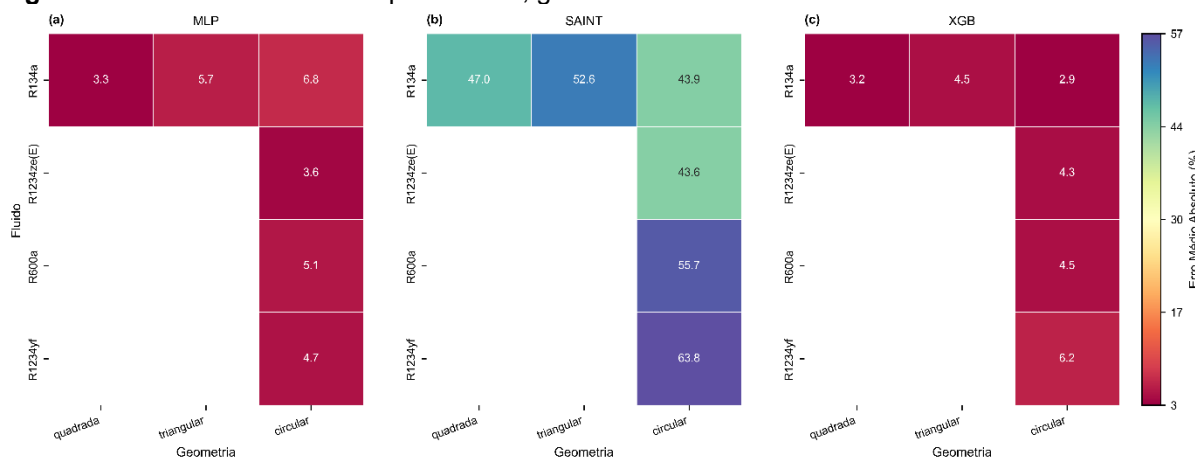
Fonte: elaborado pelo próprio autor

4.3.2 Coeficiente de troca térmica

A Figura 18 indica novamente a superioridade do XGBoost para todas as combinações apresentadas, com valores de EMA reduzidos para os quatro fluidos

refrigerantes considerados na geometria circular e para as duas geometrias adicionais do R134a.

Figura 20 - Erro médio absoluto por modelo, geometria e fluido



Fonte: elaborado pelo próprio autor

4.4 Discussão comparativa

Nas figuras 21 e 22, os pontos coloridos representam os valores experimentais medidos em laboratório para cada fluido, evidenciando diretamente a variabilidade típica do processo de ebulição convectiva em microcanais. As linhas suavizadas, por sua vez, correspondem às previsões do modelo XGBoost, obtidas a partir dos mesmos dados filtrados pelas condições operacionais descritas em cada figura. Essas curvas são refinadas por técnicas de suavização justamente para evitar oscilações artificiais e, assim, representar apenas a tendência física esperada ao longo da variação do título de vapor x . Dessa forma, a comparação entre pontos e curvas permite avaliar simultaneamente o grau de aderência do modelo aos dados experimentais e a coerência físico-termodinâmica das tendências capturadas.

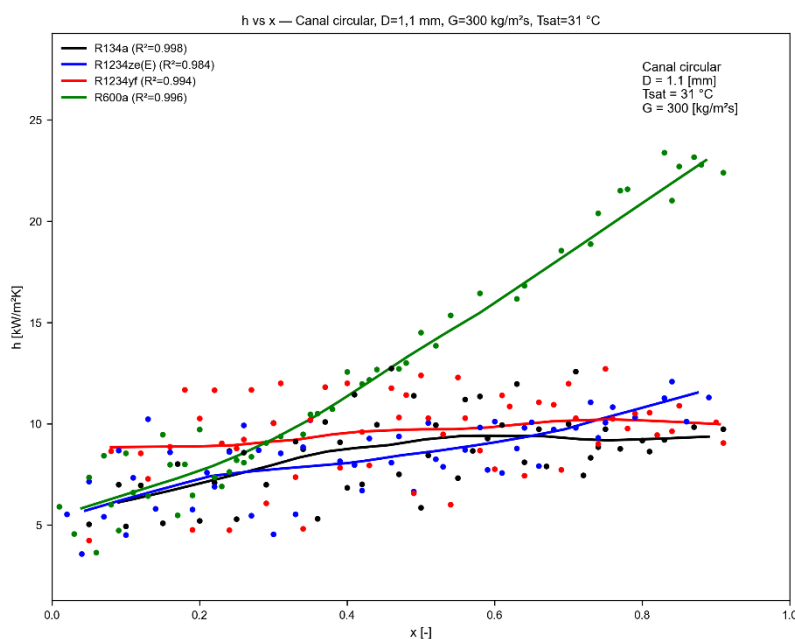
A Figura 19 apresenta a comparação direta entre as médias experimentais de h_2 e as curvas previstas pelo modelo XGBoost em função do título de vapor x , para um canal circular de diâmetro hidráulico de aproximadamente 1,1 mm sob a condição operacional $G \approx 300$ e $T_{\text{sat}} \approx 31^\circ\text{C}$. A Figura 20 apresenta a comparação direta entre as médias experimentais de dp/dz e as curvas previstas pelo modelo XGBoost em função do título de vapor x , para diâmetro de 1,1 mm, $G \approx 400$ e $T_{\text{sat}} \approx 31^\circ\text{C}$.

Para cada fluido, foram considerados apenas os pontos experimentais próximos dessa condição, o que garante que a tendência avaliada seja fisicamente consistente e referenciada a uma mesma faixa operacional.

O coeficiente de determinação R^2 indicado na legenda de cada curva foi calculado com base em todas as observações individuais filtradas por fluido, quantificando o quanto a variabilidade experimental é explicada pelo modelo naquela faixa operacional específica. Os valores elevados de R^2 obtidos em todos os fluidos demonstram que o XGBoost não apenas se ajusta bem aos pontos médios, mas também reproduz as flutuações reais presentes nos dados, o que é um aspecto desafiador devido à natureza altamente não linear da transferência de calor em escoamentos bifásicos.

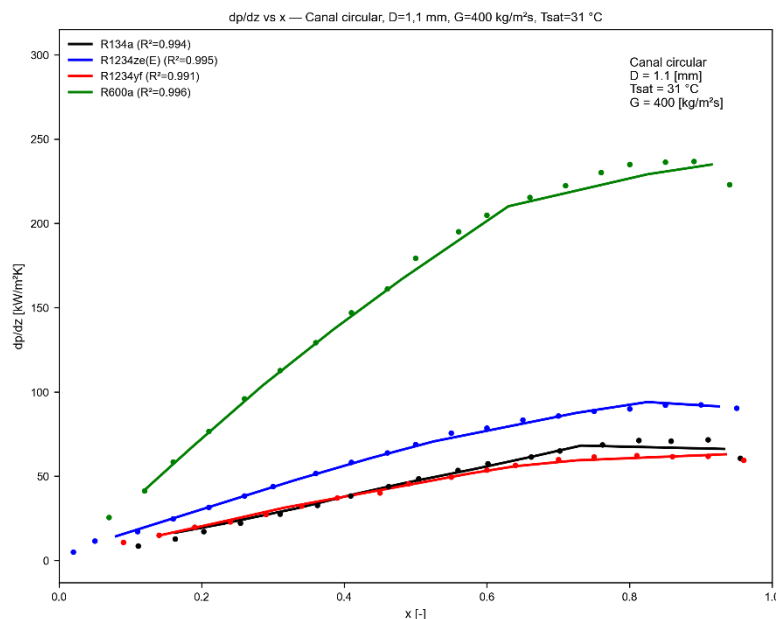
Além disso, as curvas previstas mantêm a tendência termofísica esperada ao longo do aumento do título de vapor, permanecendo dentro do envelope de incerteza experimental. Essa coerência visual reforça que o modelo captura mudanças associadas a diferentes regimes de ebulição e às propriedades termofísicas de cada fluido refrigerante de baixo GWP. Mesmo em regiões com maior dispersão experimental, a aproximação do modelo permanece consistente, sem desvios sistemáticos.

Figura 21 - Comparação entre os valores medidos e os valores previstos pelo XGBoost para o coeficiente de troca térmica



Fonte: elaborado pelo próprio autor

Figura 22 - Comparação entre os valores medidos e os valores previstos pelo XGBoost para a perda de pressão



Fonte: elaborado pelo próprio autor

5 CONCLUSÃO

Em virtude do exposto, compreende-se que este estudo investigou a aplicação de modelos de aprendizado de máquina (MLP, SAINT e XGBoost) para a predição do coeficiente de troca térmica e da perda de pressão por atrito de fluidos refrigerantes de baixo GWP em microcanais, baseando-se em dados experimentais de referência.

A escolha do tema revela-se relevante, tanto para o avanço científico no campo da termofluidodinâmica aplicada como para o setor industrial de refrigeração e eficiência térmica. Ao propor uma abordagem de aprendizado de máquina em substituição ou complemento às correlações tradicionais, o trabalho insere-se num contexto de inovação e sustentabilidade que reflete tendências atuais.

No que concerne aos objetivos inicialmente estabelecidos: i) desenvolver e avaliar os modelos de aprendizado de máquina para as variáveis termo-hidráulicas; ii) comparar quantitativamente os desempenhos dos modelos em termos de precisão, robustez e generalização; e iii) indicar o modelo mais adequado para a predição das propriedades em questão, verifica-se que todos foram cumpridos. Os três métodos foram implementados e treinados para ambas as variáveis alvo; as métricas de desempenho permitiram a comparação rigorosa entre os modelos; e a análise demonstrou que o XGBoost registrou o melhor desempenho em todas as condições avaliadas.

Em termos de resultados, o XGBoost proporcionou a menor magnitude de erro médio absoluto e as maiores taxas de predição dentro das margens de tolerância $\pm 20\%$ e $\pm 30\%$, ultrapassando a ferramenta de referência experimental. O desempenho consistente deste modelo evidencia a eficácia de algoritmos de ensemble em bases com forte não-linearidade e heterogeneidade operacional. Por sua vez, os modelos MLP e SAINT evidenciaram restrições: o MLP apresentou desempenho razoável, porém inferior ao XGBoost; o SAINT demonstrou maior variabilidade e necessidade de volume de dados superior.

Esta investigação confirma que a adoção de técnicas de aprendizado de máquina constitui uma alternativa robusta para a modelagem de fenômenos termo-hidráulicos em microcanais, contribuindo para o projeto de sistemas térmicos mais eficientes e sustentáveis. A interdisciplinaridade entre engenharia térmica, ciência de dados e termofluidodinâmica reforça a contribuição deste trabalho para a área.

REFERÊNCIAS

SEMPÉRTEGUI-TAPIA, D. F. Análise experimental do efeito da geometria da seção transversal e do desempenho de fluidos de reduzido GWP na ebulição convectiva em canais de dimensões reduzidas. 288p. Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2016.

CHEN, T.; GUESTRIN, C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. San Francisco: ACM, 2016. p. 785-794. DOI: <https://doi.org/10.1145/2939672.2939785>

BOLAJI, B. O.; HESKETH, R. P.; ADEKUNLE, A. S. Evaluation of the performance of R600a and R290 as replacements for R134a in domestic refrigerators. Applied Thermal Engineering

LOMBARDO, G. et al. Performance of low-GWP refrigerants R1234yf and R1234ze(E) in small-scale vapor compression systems. International Journal of Refrigeration

SOMEPALI, G. et al. SAINT: Improved neural networks for tabular data via row attention and contrastive pre-training. arXiv preprint, arXiv:2106.01342, 2021. Disponível em: <https://arxiv.org/abs/2106.01342>

ZHAO, Y. et al. Liquid viscosity and surface tension of R1234yf and R1234ze(E) under saturation conditions by surface light scattering. Journal of Chemical & Engineering Data, v. 59, n. 10, p. 3123–3131, 2014.

LEMMON, E. W. et al. NIST Chemistry WebBook, NIST Standard Reference Database Number 69. National Institute of Standards and Technology, 2025.

BELL, I. H. et al. Pure and pseudo-pure fluid thermophysical property evaluation and the open-source thermophysical property library CoolProp. Industrial & Engineering Chemistry Research, v. 53, n. 6, p. 2498–2508, 2014.

ZOU, J. et al. Recent advances in the applications of machine learning methods for heat exchanger modeling—a review. Frontiers in Energy Research, v. 11, p. 1–19, 2023. DOI: <https://doi.org/10.3389/fenrg.2023.1294531>.

CANDAL, R. J. et al. Two-phase flow regimes in small channels. Experimental Thermal and Fluid Science, v. 88, p. 180–192, 2017.

SOUZA, F. N. et al. Aplicações de redes neurais artificiais em processos industriais. Revista Engenharia e Tecnologia Aplicada, v. 7, p. 45–58, 2012.

