

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM
COMPUTACIONAL

Rômulo Murucci Oliveira

**Predição da resistência à compressão simples de solos estabilizados utilizando
aprendizado de máquina automatizado**

Juiz de Fora

2026

Rômulo Murucci Oliveira

**Predição da resistência à compressão simples de solos estabilizados utilizando
aprendizado de máquina automatizado**

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Modelagem Computacional. Área de concentração: Modelagem Computacional.

Orientador: Prof. Dr. Leonardo Goliatt da Fonseca

Coorientadora: Prof^a. Dr^a. Katia Vanessa Bicalho

Juiz de Fora

2026

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF
com os dados fornecidos pelo(a) autor(a)

Oliveira, Rômulo Murucci.

Predição da resistência à compressão simples de solos estabilizados utilizando aprendizado de máquina automatizado / Rômulo Murucci Oliveira. – 2026.

116 f. : il.

Orientador: Leonardo Goliatt da Fonseca

Coorientadora: Katia Vanessa Bicalho

Dissertação (Mestrado) – Universidade Federal de Juiz de Fora, Faculdade de Engenharia. Programa de Pós-Graduação em Modelagem Computacional, 2026.

1. autoML. 2. resistência à compressão simples. 3. solos estabilizados. 4. predição. 5. aprendizado de máquina automatizado. I. da Fonseca, Leonardo Goliatt, orient. II. Bicalho, Katia Vanessa, coorient. III. Título.

Rômulo Murucci Oliveira

Predição da resistência à compressão simples de solos estabilizados utilizando aprendizado de máquina automatizado

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Modelagem Computacional. Área de concentração: Modelagem Computacional.

Aprovada em 9 de março de 2026.

BANCA EXAMINADORA

Prof. Dr. Leonardo Goliatt da Fonseca - Orientador

Universidade Federal de Juiz de Fora

Prof. Dr. Wanderlei Malaquias Pereira Junior

Universidade Federal de Catalão

Prof^a. Dr^a. Tatiana Tavares Rodriguez

Universidade Federal de Juiz de Fora

Juiz de Fora, 14/04/2026.



Documento assinado eletronicamente por **Leonardo Goliatt da Fonseca, Professor(a)**, em 14/04/2026, às 07:54, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Wanderlei Malaquias Pereira Junior, Usuário Externo**, em 14/04/2026, às 11:08, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Tatiana Tavares Rodriguez, Professor(a)**, em 14/04/2026, às 19:09, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **2946822** e o código CRC **09D8D650**.

AGRADECIMENTOS

Primordialmente, agradeço a **Deus** por me conceder força e sabedoria para trilhar este caminho e persistir na busca pelos meus sonhos. Em todos os momentos, especialmente naqueles em que a pressão e o estresse se fizeram presentes, agradeço pelo amor, carinho e apoio incondicional da minha esposa, **Rebeca**; sem o seu incentivo, não teria sido possível superar este desafio. À minha **família e amigos**, agradeço pela paciência e compreensão diante das minhas ausências durante a dedicação aos estudos. Manifesto um agradecimento especial ao meu amigo **João Vitor**, pelos valiosos conselhos e contribuições na resolução de problemas técnicos relacionados ao desenvolvimento desta dissertação.

Expresso minha gratidão à **Universidade Federal de Juiz de Fora (UFJF)** e ao **Programa de Pós-Graduação em Modelagem Computacional (PGMC)**. Agradeço ao corpo docente, aos(às) funcionários(as), à direção e à coordenação, que proporcionaram as condições necessárias para o meu desenvolvimento profissional, crescimento pessoal e para a capacidade de resolução de problemas. Ao meu orientador, **Leonardo**, agradeço por aceitar este desafio e por todo o suporte fundamental; sua orientação foi essencial para a viabilização deste trabalho. À minha coorientadora, **Katia**, agradeço pela paciência e pelos valiosos conselhos técnicos. Aos(Às) colegas do PGMC, membros e egressos, meu obrigado pelos anos de aprendizado, trabalho em equipe e experiências inesquecíveis.

Agradeço aos(às) colegas da Coordenação de Programação de Trens da **MRS Logística**, especialmente ao Adalberto, Alex, Henrique, Leonam, Patrick e Stefson, pela compreensão e pelo apoio nas trocas de escalas, permitindo minha frequência às aulas. Estendo este agradecimento aos meus antigos gestores, Fernando, Letícia e Rodrigo, pelo suporte na flexibilização de horários e no incentivo à participação em eventos e congressos acadêmicos.

A todas as pessoas que, direta ou indiretamente, contribuíram para a minha formação, manifesto o meu mais sincero agradecimento.

RESUMO

Estabilização de solos com aditivos de vários tipos, mais comumente aditivos químico do tipo cal e cimento, têm sido usados para melhorar as propriedades dos solos em obras geotécnicas por séculos. Resultados do ensaio de resistência à compressão simples (RCS) são utilizados para controle e avaliação do aumento da resistência ao cisalhamento de solos estabilizados. Entretanto, a grande quantidade de ensaios necessários em amostras representativas de solo estabilizado torna as avaliações experimentais demoradas e dispendiosas. O objetivo deste trabalho é aplicar e comparar cinco sistemas de aprendizado de máquina automatizado (AutoML) — AutoGluon, AutoKeras, FLAML, H2O e TPOT — na predição da RCS de solos estabilizados. Diferente de abordagens anteriores que automatizavam apenas o ajuste de hiperparâmetros, esta pesquisa propõe um fluxo de automatização integral, incluindo pré-processamento, seleção de atributos e definição algorítmica. A metodologia consistiu na análise de dez conjuntos de dados experimentais com variadas condições geotécnicas, totalizando 1.745 amostras provenientes de diversos países e abrangendo estabilizantes como cimento, cal e resíduos agrícolas. O desempenho foi avaliado por meio de métricas de regressão e um Índice de Desempenho (ID) composto. Os resultados indicam que os sistemas de AutoML superaram consistentemente os modelos reportados na literatura original de cada base de dados. O sistema FLAML apresentou o desempenho global mais estável e eficiente, com o maior ID médio (0,78). A análise de importância dos atributos revelou que os modelos capturaram com sucesso mecanismos físicos e químicos fundamentais, priorizando o teor de estabilizantes e parâmetros de compactação. Conclui-se que o AutoML é uma ferramenta robusta para a engenharia geotécnica, permitindo a construção rápida de modelos preditivos e reduzindo a dependência de processos experimentais extensos, além de mitigar a necessidade de conhecimento especializado em ciência de dados, embora modelos manuais especificamente otimizados ainda possam atingir precisões marginais superiores.

Palavras-chave: autoML; resistência à compressão simples; solos estabilizados; predição; aprendizado de máquina automatizado.

ABSTRACT

Soil stabilization using various types of additives, most commonly chemical additives such as lime and cement, has been used for centuries to improve soil properties in geotechnical engineering projects. Results from the unconfined compressive strength (UCS) test are used to monitor and evaluate the increase in shear strength of stabilized soils. However, the large number of tests required on representative samples of stabilized soil makes experimental evaluations time-consuming and costly. The objective of this study is to apply and compare five automated machine learning (AutoML) systems—AutoGluon, AutoKeras, FLAML, H2O, and TPOT—for predicting the UCS of stabilized soils. Unlike previous approaches that automated only hyperparameter tuning, this research proposes a fully automated workflow, including preprocessing, feature selection, and algorithm selection. The methodology consisted of analyzing ten experimental datasets under varied geotechnical conditions, totaling 1,745 samples from multiple countries and covering stabilizing agents such as cement, lime, and agricultural residues. Performance was evaluated using regression metrics and a composite Performance Index (ID). The results indicate that the AutoML systems consistently outperformed the models reported in the original literature for each dataset. The FLAML system demonstrated the most stable and efficient overall performance, achieving the highest average ID (0.78). Feature importance analysis revealed that the models successfully captured fundamental physical and chemical mechanisms, prioritizing stabilizer content and compaction parameters. It is concluded that AutoML is a robust tool for geotechnical engineering, enabling the rapid development of predictive models and reducing reliance on extensive experimental procedures, while also mitigating the need for specialized expertise in data science. Nevertheless, specifically optimized manual models may still achieve marginally higher levels of accuracy.

Keywords: autoML; unconfined compressive strength; stabilized soils; prediction; automated machine learning.

LISTA DE ILUSTRAÇÕES

Figura 1 – Níveis de automação nas principais etapas da modelagem de aprendizado de máquina em estudos prévios de predição de RCS, comparados ao fluxo AutoML integral empregado neste trabalho.	17
Figura 2 – Equipamento para realização do ensaio de compressão simples.	20
Figura 3 – Ensaio de compressão simples em corpos de prova de argila, à esquerda argila pura, e à direita argila estabilizada com biopolímero.	21
Figura 4 – Fluxograma das etapas do aprendizado de máquina.	24
Figura 5 – Resumo de grupos de problemas em AutoML, com exemplos de espaço de busca, estratégias de busca e formas de avaliação.	29
Figura 6 – Denominações das abordagens de AutoML para a geração e avaliação de modelos.	30
Figura 7 – Fluxograma esquemático do AutoGluon.	30
Figura 8 – Fluxograma esquemático do AutoKeras.	32
Figura 9 – Fluxograma esquemático do FLAML.	32
Figura 10 – Fluxograma esquemático do H2O.	33
Figura 11 – Fluxograma esquemático do TPOT.	34
Figura 12 – Classificação dos solos de acordo com o USCS.	36
Figura 13 – Distribuição geográfica dos locais das amostras dos conjuntos de dados.	36
Figura 14 – Histogramas das variáveis do conjunto de dados D1.	38
Figura 15 – Histogramas das variáveis do conjunto de dados D2.	39
Figura 16 – Histogramas das variáveis do conjunto de dados D3.	41
Figura 17 – Histogramas das variáveis do conjunto de dados D4.	42
Figura 18 – Histogramas das variáveis do conjunto de dados D5.	44
Figura 19 – Histogramas das variáveis do conjunto de dados D6.	46
Figura 20 – Histogramas das variáveis do conjunto de dados D7.	47
Figura 21 – Histogramas das variáveis do conjunto de dados D8.	49
Figura 22 – Histogramas das variáveis do conjunto de dados D9.	50
Figura 23 – Histogramas das variáveis do conjunto de dados D10.	52
Figura 24 – Informações disponíveis nos conjuntos de dados analisados.	53
Figura 25 – Exemplo de validação cruzada <i>k-fold</i> com cinco subconjuntos (<i>folds</i>).	56
Figura 26 – Siglas e expressões das métricas de desempenho e do desvio padrão.	57
Figura 27 – Tempo médio de execução para cada conjunto de dados e sistema AutoML.	61
Figura 28 – Distribuição dos resultados por sistemas para a métrica R.	62
Figura 29 – Distribuição dos resultados por sistemas para a métrica R ²	62
Figura 30 – Distribuição dos resultados por sistemas para a métrica RMSE.	63
Figura 31 – Distribuição dos resultados por sistemas para a métrica MAE.	63

Figura 32 – Distribuição dos resultados por sistemas para a métrica MAPE.	64
Figura 33 – Melhores modelos utilizados.	69
Figura 34 – Diferenças significativas do teste de Kruskal-Wallis entre sistemas.	72
Figura 35 – Diferenças significativas entre pares do teste de Dunn entre sistemas.	73
Figura 36 – Mapa de calor da quantidade de diferenças significativas entre o AutoGluon e os demais sistemas.	75
Figura 37 – Mapa de calor da quantidade de diferenças significativas entre o AutoKeras e os demais sistemas.	75
Figura 38 – Mapa de calor da quantidade de diferenças significativas entre o FLAML e os demais sistemas.	76
Figura 39 – Mapa de calor da quantidade de diferenças significativas entre o H2O e os demais sistemas.	76
Figura 40 – Mapa de calor da quantidade de diferenças significativas entre o TPOT e os demais sistemas.	77
Figura 41 – Importância média dos atributos para o conjunto de dados D1.	78
Figura 42 – Importância média dos atributos para o conjunto de dados D2.	79
Figura 43 – Importância média dos atributos para o conjunto de dados D3.	80
Figura 44 – Importância média dos atributos para o conjunto de dados D4.	81
Figura 45 – Importância média dos atributos para o conjunto de dados D5.	82
Figura 46 – Importância média dos atributos para o conjunto de dados D6.	83
Figura 47 – Importância média dos atributos para o conjunto de dados D7.	84
Figura 48 – Importância média dos atributos para o conjunto de dados D8.	85
Figura 49 – Importância média dos atributos para o conjunto de dados D9.	86
Figura 50 – Importância média dos atributos para o conjunto de dados D10.	87

LISTA DE TABELAS

Tabela 1 – Comparativo dos sistemas de AutoML.	31
Tabela 2 – Resumo dos conjuntos de dados.	35
Tabela 3 – Estatísticas básicas do conjunto de dados D1.	37
Tabela 4 – Estatísticas básicas do conjunto de dados D2.	38
Tabela 5 – Estatísticas básicas do conjunto de dados D3.	40
Tabela 6 – Estatísticas básicas do conjunto de dados D4.	42
Tabela 7 – Estatísticas básicas do conjunto de dados D5.	43
Tabela 8 – Propriedades físicas e classificação dos solos do conjunto de dados D5.	44
Tabela 9 – Estatísticas básicas do conjunto de dados D6.	45
Tabela 10 – Estatísticas básicas do conjunto de dados D7.	46
Tabela 11 – Propriedades físicas e classificação dos solos do conjunto de dados D7.	48
Tabela 12 – Estatísticas básicas do conjunto de dados D8.	49
Tabela 13 – Estatísticas básicas do conjunto de dados D9.	50
Tabela 14 – Estatísticas básicas do conjunto de dados D10.	51
Tabela 15 – Hiperparâmetros definidos para cada sistema de AutoML.	54
Tabela 16 – Bibliotecas e versões utilizadas.	60
Tabela 17 – Valores médios das métricas de desempenho. (-) indica valores não disponíveis.	65
Tabela 18 – Valores médios das métricas de desempenho. (-) indica valores não disponíveis (continuação).	66
Tabela 19 – Desvios padrão das métricas de desempenho.	67
Tabela 20 – Desvios padrão das métricas de desempenho (continuação).	68
Tabela 21 – Comparação binária de desempenho entre os sistemas AutoML e os modelos de referência.	70
Tabela 22 – Desempenho médio (ID) e desvio padrão dos sistemas AutoML para cada conjunto de dados.	71
Tabela 23 – Resumo dos melhores sistemas AutoML por conjunto de dados.	71
Tabela 24 – Quantidade de diferenças significativas entre sistemas por métrica, conforme o teste de Dunn.	74
Tabela 25 – Quantidade de diferenças significativas entre sistemas por conjuntos de dados, conforme o teste de Dunn.	74
Tabela 26 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D1.	107
Tabela 27 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D2.	108

Tabela 28 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D3.	109
Tabela 29 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D4.	110
Tabela 30 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D5.	111
Tabela 31 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D6.	112
Tabela 32 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D7.	113
Tabela 33 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D8.	114
Tabela 34 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D9.	115
Tabela 35 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D10.	116

LISTA DE ABREVIATURAS E SIGLAS

AG	AutoGluon
AK	AutoKeras
ANN	<i>Artificial Neural Networks</i>
ANOVA	Análise de variância
ASTM	<i>American Society for Testing and Materials</i>
AutoML	<i>Automated Machine Learning</i>
BRT	<i>Boosted Regression Tree</i>
CLR	mistura de cimento, cal e cinza de casca de arroz
DP	Desvio padrão
DRF	<i>Distributed Random Forest</i>
DT	<i>Decision Tree</i>
EUA	Estados Unidos da América
FL	FLAML
FLAML	<i>Fast Lightweight Automated Machine Learning</i>
GA	<i>Genetic Algorithms</i>
GBM	<i>Gradient Boosting Machine</i>
GFRP	<i>Glass fiber-reinforced polymer</i>
GGBS	<i>Ground granulated blast furnace slag</i>
GLM	<i>Generalized Linear Model</i>
GPR	<i>Gaussian Process Regression</i>
GWO	<i>Grey Wolf Optimizer</i>
H2	H2O
ICA	<i>Independent Component Analysis</i>
ID	Índice de Desempenho
LDA	<i>Linear Discriminant Analysis</i>
MAE	<i>Mean Absolute Error</i>
MAPE	<i>Mean Absolute Percentage Error</i>
MEP	<i>Multi Expression Programming</i>
MICP	<i>Microbially induced calcite precipitation</i>
ML	<i>Machine Learning</i>
MLR	<i>Multiple Linear Regression</i>
PCA	<i>Principal Component Analysis</i>
pH	Potencial hidrogeniônico
PSO	<i>Particle Swarm Optimization</i>
RCS	Resistência à compressão simples
Ref.	Referência
Resist.	Resistência
RF	<i>Random Forest</i>
RMSE	<i>Root Mean Squared Error</i>
SHAP	<i>Shapley Additive Explanations</i>

SVM	<i>Support Vector Machines</i>
SVR	<i>Support Vector Regression</i>
TP	TPOT
TPOT	<i>Tree-Based Pipeline Optimization Tool</i>
USCS	<i>Unified Soil Classification System</i>
XGB	<i>Extreme Gradient Boosting</i>
XRT	<i>Extremely Randomized Trees</i>

SUMÁRIO

1	INTRODUÇÃO	14
1.1	MOTIVAÇÃO	15
1.2	OBJETIVO	18
1.3	ESTRUTURA DO TRABALHO	18
2	REVISÃO BIBLIOGRÁFICA	19
2.1	ESTABILIZAÇÃO DE SOLOS	19
2.1.1	ENSAIO DE COMPRESSÃO SIMPLES	19
2.1.2	ESTABILIZAÇÃO QUÍMICA	21
2.1.3	ESTABILIZAÇÃO POR COMPACTAÇÃO E POR REFORÇO	23
2.2	APRENDIZADO DE MÁQUINA	23
2.2.1	MANIPULAÇÃO DOS DADOS	24
2.2.2	ENGENHARIA DE ATRIBUTOS	26
2.2.3	GERAÇÃO DO MODELO	26
2.2.4	AVALIAÇÃO DO MODELO	27
2.3	SISTEMA DE APRENDIZADO DE MÁQUINA AUTOMATIZADO	28
2.3.1	AUTOGLUON	29
2.3.2	AUTOKERAS	31
2.3.3	<i>FAST LIGHTWEIGHT AUTOML (FLAML)</i>	31
2.3.4	H2O	32
2.3.5	<i>TREE-BASED PIPELINE OPTIMIZATION TOOL (TPOT)</i>	33
3	MATERIAIS E MÉTODOS	35
3.1	CONJUNTOS DE DADOS	35
3.1.1	CONJUNTOS DE DADOS D1 E D2	37
3.1.2	CONJUNTO DE DADOS D3	40
3.1.3	CONJUNTO DE DADOS D4	41
3.1.4	CONJUNTO DE DADOS D5	43
3.1.5	CONJUNTO DE DADOS D6	45
3.1.6	CONJUNTO DE DADOS D7	46
3.1.7	CONJUNTO DE DADOS D8	48
3.1.8	CONJUNTO DE DADOS D9	50
3.1.9	CONJUNTO DE DADOS D10	51
3.1.10	COMPARAÇÃO DOS CONJUNTOS DE DADOS	52
3.2	DESENVOLVIMENTO E VALIDAÇÃO DOS SISTEMAS	53
3.3	MÉTRICAS DE DESEMPENHO	57
3.4	IMPORTÂNCIA DOS ATRIBUTOS	58
3.5	CONFIGURAÇÕES COMPUTACIONAIS	59
4	RESULTADOS	61

4.1	MÉTRICAS DE DESEMPENHO	61
4.1.1	ANÁLISE DESCRITIVA	68
4.1.2	ANÁLISE ESTATÍSTICA	72
4.2	IMPORTÂNCIA DOS ATRIBUTOS	77
4.3	INTERPRETAÇÕES DOS RESULTADOS	88
5	CONCLUSÕES	91
5.1	LIMITAÇÕES	92
5.2	TRABALHOS FUTUROS	93
	REFERÊNCIAS	95
	APÊNDICE A – Testes de normalidade e homogeneidade . .	107

1 INTRODUÇÃO

A obtenção de valores precisos e confiáveis de resistência à compressão simples (RCS) é essencial para o manejo eficiente e sustentável de solos estabilizados. Essa propriedade está diretamente relacionada às características físicas do solo, ao traço adotado na mistura estabilizadora e aos procedimentos aplicados durante a determinação experimental dos resultados (1). Além disso, cresce o interesse no uso de resíduos regionais de fácil acesso como substitutos parciais do cimento, visando reduzir custos e impactos ambientais. Melhorias nas propriedades físicas, químicas, mecânicas e elétricas dos solos estabilizados também têm se mostrado relevantes para o aprimoramento do desempenho desses materiais (2).

A estabilização de solos naturais mostra-se uma alternativa sustentável para a construção civil, pois reduz custos e impactos ambientais ao modificar as propriedades físicas com o objetivo de obter melhor comportamento mecânico (adicionar referência). Entre as técnicas disponíveis, a estabilização química é amplamente utilizada, com a adição de materiais como cal e cimento que reagem com os minerais do solo, promovendo melhorias significativas em seu comportamento (3, 4). Apesar do cimento Portland e a cal serem os agentes ligantes mais comuns, seu uso contínuo está associado a elevados impactos ambientais, sobretudo pela emissão de dióxido de carbono (5). Portanto, estudos recentes investigam o aproveitamento de resíduos agrícolas como alternativa, podendo atuar como aditivo ou substituto parcial do cimento em misturas de solos estabilizados (6, 2). A determinação de propriedades importantes, como a resistência à compressão simples, em solos estabilizados com diferentes combinações de ligantes, exige um conjunto de ensaios experimentais que envolve a definição das proporções de solo e aglutinante, o tempo de cura e a análise de parâmetros que influenciam o comportamento final do material.

Nos últimos anos, métodos baseados em aprendizado de máquina (*Machine Learning* - ML) têm se mostrado uma alternativa promissora para superar as limitações dos procedimentos tradicionais utilizados na predição da resistência à compressão simples de solos (7, 8). Os ensaios laboratoriais convencionais exigem grande esforço experimental, consumo de tempo e elevado custo para a obtenção de resultados confiáveis. Além disso, técnicas estatísticas clássicas, como a análise de regressão, muitas vezes se mostraram insuficientes, pois são baseadas em misturas não representativas e não conseguem capturar relações complexas entre as variáveis envolvidas. Nesse contexto, o uso de algoritmos de aprendizado de máquina tem permitido analisar bases de dados experimentais e identificar padrões ocultos de comportamento. Entre os modelos avaliados, a Floresta Aleatória (*Random Forest* - RF) tem apresentado bom desempenho preditivo e capacidade de generalização superior quando comparada a técnicas como Regressão de Vetor Suporte (*Support Vector Regression* - SVR) e Redes Neurais Artificiais (*Artificial Neural Networks* - ANN). O

RF também se destaca por sua resistência a ruídos e valores discrepantes, além de reduzir o risco de sobreajuste. No entanto, seu desempenho pode ser afetado pela sensibilidade aos hiperparâmetros, o que tem motivado o uso de algoritmos de otimização em substituição ao ajuste manual tradicional (9).

Diversos estudos têm aplicado técnicas de aprendizado de máquina na predição da RCS de solos, com destaque para os modelos baseados em Árvores de Decisão (*Decision Trees* - DT) e RF, reconhecidos por apresentarem boa precisão e interpretação direta (10, 11, 12, 13, 14, 15). Máquinas de Vetores de Suporte (*Support Vector Machines* - SVM) também têm sido amplamente utilizadas pela sua capacidade de lidar com relações não lineares entre as variáveis (16, 17, 18, 12). Em paralelo, a Programação Genética (*Genetic Algorithms* - GA) tem sido usada para derivar expressões explícitas entre a RCS e as propriedades dos solos (19, 20, 21, 22). Pesquisas mais recentes destacam os métodos de aprendizado por conjunto, como o *Gradient Boosting Machine* (GBM), que combina diversos modelos simples para formar predições mais precisas, mostrando bom desempenho mesmo em dados com ruído e não linearidade (1, 23, 24, 25, 26).

O aprimoramento desses modelos ocorreu por meio de algoritmos de otimização de hiperparâmetros, como Otimização por Enxame de Partículas (*Particle Swarm Optimization* - PSO), GA e o *Bat Algorithm* (27, 28, 29, 30, 31). As ANNs, contudo, continuam sendo as mais utilizadas na predição da RCS, dada sua flexibilidade e boa capacidade de representar comportamentos complexos (10, 32, 11, 33), sendo aplicadas em diferentes contextos geotécnicos, como solos reforçados com fibras (34, 35, 36), fundações de pavimentos (37, 38, 39), estabilização de solos com cimento (40) e processos construtivos de terraplenagem (41, 42). Esses avanços reforçam o potencial das abordagens baseadas em dados como uma alternativa eficiente para reduzir o número de ensaios laboratoriais e apoiar o projeto de estabilização de solos.

1.1 MOTIVAÇÃO

A estabilização de solos com o uso de materiais como cimento, cal e cinza volante é uma prática consolidada na engenharia geotécnica, com o objetivo de melhorar o desempenho mecânico e a durabilidade dos solos em obras de infraestrutura (43, 44). A definição das proporções ideais desses aditivos ainda depende fortemente de procedimentos experimentais, que consistem na preparação de amostras com diferentes teores de estabilizante, seguidos de ensaios de compressão simples (45, 46, 47). Esse processo é repetido até que se obtenha a resistência desejada, podendo demandar várias semanas devido ao tempo necessário para a moldagem, cura e ensaio das amostras (48).

Esse método apresenta limitações práticas, especialmente quando há necessidade de avaliar diversas combinações de materiais ou quando o tempo disponível para o projeto é reduzido. Nesse contexto, o uso de técnicas de aprendizado de máquina tem se mostrado

uma alternativa promissora, pois permite estimar valores de RCS a partir de propriedades básicas do solo e de variáveis relacionadas aos estabilizantes. Essa abordagem pode reduzir o número de ensaios de laboratório e auxiliar no planejamento de misturas mais eficientes (48, 49, 50, 51, 3, 52, 53, 20, 54).

Mais recentemente, o aprendizado de máquina automatizado (*Automated Machine Learning* – AutoML) tem se destacado por permitir a construção de modelos de forma mais rápida e sistemática. Essa abordagem automatiza diversas etapas do processo tradicional de modelagem de dados (ou *pipeline* de dados), como a limpeza, seleção de variáveis, escolha do algoritmo, ajuste de parâmetros e avaliação de desempenho, podendo ser aplicada por usuários com diferentes níveis de experiência (55, 56, 57). Atualmente, as técnicas de AutoML atingiram um grau de maturidade que lhes permite rivalizar e, em alguns casos, superar o desempenho de especialistas em aprendizado de máquina (57).

Uma busca por trabalhos que aplicam AutoML na predição da RCS de solos foi realizada nas bases *Scopus* e *Web of Science*, em 08 de junho de 2025. Foram utilizadas as expressões (*unconfined compressive strength*) AND (*soil*) AND (*automated machine learning OR AutoML OR auto machine learning*), e apenas dois estudos relevantes foram identificados (ambos de 2025). Nos dois casos, a automatização foi restrita à etapa de otimização de hiperparâmetros, sem a automatização completa do fluxo de aprendizado de máquina. No estudo de Goliatt et al. (58) foi utilizado o algoritmo evolutivo *Grey Wolf Optimizer* (GWO) para selecionar as variáveis mais relevantes no treinamento de um modelo de GBM. Já Wang (59) comparou três algoritmos evolutivos (*Tunicate Swarm Algorithm*, *Sea-horse Optimizer* e DT) para otimizar a seleção de variáveis na predição da RCS.

Dessa forma, utilizou-se a busca de trabalhos nas bases *Scopus* e *Web of Science* em 08 de junho de 2025, utilizando as expressões (*unconfined compressive strength*) AND (*soil*) AND (*machine learning OR artificial intelligence OR soft computing*), com foco nos trabalhos mais recentes e suas respectivas referências.

Não foram encontrados trabalhos que aplicassem AutoML de forma ampla na predição da RCS, apenas em alguma etapa ou técnica na modelagem do aprendizado de máquina. Essa ausência representa uma lacuna na literatura e uma oportunidade de pesquisa, que este estudo busca investigar e aprofundar.

Grande parte dos estudos que combinam métodos de otimização com algoritmos de aprendizagem de máquina para estimar a RCS concentra-se apenas na etapa de ajuste de hiperparâmetros. Nessas abordagens, tarefas essenciais (como preparação dos dados, seleção de atributos e escolha dos modelos) continuam altamente dependentes de intervenção manual, o que limita a escalabilidade das soluções e prejudica a reprodutibilidade dos resultados.

O fluxo de trabalho proposto neste estudo adota uma perspectiva mais ampla: em

vez de automatizar apenas o ajuste fino dos modelos, incorpora mecanismos automáticos para todas as etapas fundamentais da modelagem de aprendizado de máquina. Isso inclui desde o pré-processamento dos dados até a seleção de características, a definição do algoritmo mais adequado e a otimização dos hiperparâmetros. Essa automatização integral reduz a influência de decisões subjetivas, diminui a necessidade de conhecimento especializado e aprimora a consistência entre diferentes análises (60). A Figura 1 apresenta um panorama comparativo entre trabalhos representativos da literatura, que combinam ML tradicional com métodos meta-heurísticos, e o fluxo AutoML completo adotado no presente estudo.

Figura 1 – Níveis de automação nas principais etapas da modelagem de aprendizado de máquina em estudos prévios de predição de RCS, comparados ao fluxo AutoML integral empregado neste trabalho.

Estudo	Pré processamento	Seleção de atributos	Definição do modelo	Ajuste de hiperparâmetros	Avaliação	Tipo de <i>Workflow</i>
Kardani et al. (2021)	Manual	Não aplicam	Modelos regressivos pré-definidos	Algoritmo evolutivo	Manual	Semi-automatizado
Ngo et al. (2023)	Manual	Verificações manuais de correlação	Modelo baseado em SVM previamente escolhido	Otimização via PSO/GA	Manual	Semi-automatizado
Onyelowe et al. (2024)	Entradas padronizadas	Filtragem estatística	Comparação entre modelos do tipo <i>ensemble</i>	Busca em grade/aleatória	Validação cruzada manual	Semi-automatizado
Yao et al. (2024)	Manual	Priorização de atributos via SHAP	Arquitetura híbrida baseada em XGB	Otimização híbrida baseada no XGB	Manual	Semi-automatizado
Teodoru et al. (2025)	Limpeza automatizada	Eliminação recursiva de atributos	Seleção de modelos com Auto-sklearn	Otimização Bayesiana (Optuna)	Automática	Semi-automatizado
Este estudo	Automático	Automática	Automática	Automático	Automática	Automatização integral

Fonte: Elaborado pelo autor (2025).

1.2 OBJETIVO

A pesquisa contempla a análise de cinco sistemas AutoML de destaque, disponíveis na literatura e em plataformas de código aberto. Esses sistemas serão aplicados em bases de dados provenientes de diferentes países, o que permite avaliar a capacidade dos modelos em lidar com amostras heterogêneas e condições geotécnicas variadas. Essa abordagem busca contribuir para a construção de modelos mais robustos e generalizáveis, capazes de representar com maior fidelidade o comportamento de misturas solo-estabilizante.

A pesquisa desenvolvida nesta dissertação teve como objetivo geral realizar a predição da RCS de solos estabilizados, a partir de propriedades básicas do solo e de informações sobre os estabilizantes, utilizando AutoML. Com o propósito de se alcançar o objetivo geral supracitado, foram definidos os seguintes objetivos específicos: (i) automatizar, de forma integral, todas as etapas do aprendizado de máquina (ii) comparar diferentes sistemas de AutoML em termos de precisão preditiva e eficiência; (iii) discutir o potencial do AutoML como ferramenta de apoio à engenharia geotécnica, reduzindo a dependência de processos experimentais demorados e oferecendo subsídios para a tomada de decisão em projetos de estabilização de solos; (iv) mitigar a necessidade de conhecimento especializado em ciência de dados para testar os modelos de aprendizado de máquina.

1.3 ESTRUTURA DO TRABALHO

Esta dissertação está organizada da seguinte forma: o Capítulo 2 consolida os principais conceitos da mecânica de solos relacionados à estabilização de solos, e detalha as etapas do aprendizado de máquina e os sistemas de aprendizado de máquina automatizado utilizados neste estudo. Em seguida, o Capítulo 3 apresenta as bases de dados utilizadas e descreve o arcabouço computacional adotado. O Capítulo 4 traz os resultados e a discussão dos experimentos realizados. Por fim, o Capítulo 5 apresenta as conclusões e sugestões para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

Este Capítulo revisa a literatura científica acerca dos temas de estabilização de solos e de aprendizado de máquina automatizado.

2.1 ESTABILIZAÇÃO DE SOLOS

De modo geral, o tratamento de solos pode ocorrer de duas formas distintas. A primeira consiste na modificação das propriedades do solo por meio da adição de ligante, caracterizando o chamado melhoramento do solo. Já quando essa adição promove ganhos significativos no comportamento mecânico do material, configura-se o processo de estabilização (66). O presente trabalho aborda solos estabilizados para utilização como material de construção.

2.1.1 ENSAIO DE COMPRESSÃO SIMPLES

Nos cálculos de engenharia civil geotécnica, o estudo de relações entre tensões e deformações é necessário para avaliar questões de resistência ao cisalhamento de solos naturais ou solos estabilizados como por exemplo: capacidade de carga de fundações e estabilidade de taludes. Em muitos problemas de solos o mais importante é a deformação, e não a tensão que provocará sua ruptura. A tensão necessária para que um solo se rompa sob uma fundação superficial (por exemplo sapata) implica em deformações tão grandes que a estrutura, para acompanhar tais deformações, se romperá bem antes do solo. No caso de uma rocha intacta, a ruptura é evidente e visível pela fragmentação da rocha, já no caso de uma argila mole a ruptura é uma questão de definição por meio de uma deformação considerada excessiva. Na maioria dos solos, a tensão de ruptura costuma ser definida como a tensão máxima suportada pelo solo, quer haja fragmentação ou não do mesmo (67, 68, 69).

A resistência ao cisalhamento de solos depende de vários fatores (por exemplo: composição mineralógica do solo, micro e macroestrutura, história de tensão), e uma formulação que inclua todos esses fatores não é definida na área de geotecnia (67, 70, 71, 72). Os valores de resistência ao cisalhamento de solos são determinados por meio de ensaios específicos. Nesta dissertação, avalia-se a predição numérica de resultados experimentais obtidos em laboratório, referentes à resistência ao cisalhamento de solos e rochas, a qual é estimada a partir da resistência à compressão simples, obtida por meio de ensaios de compressão não confinada (também chamados de ensaios de compressão simples).

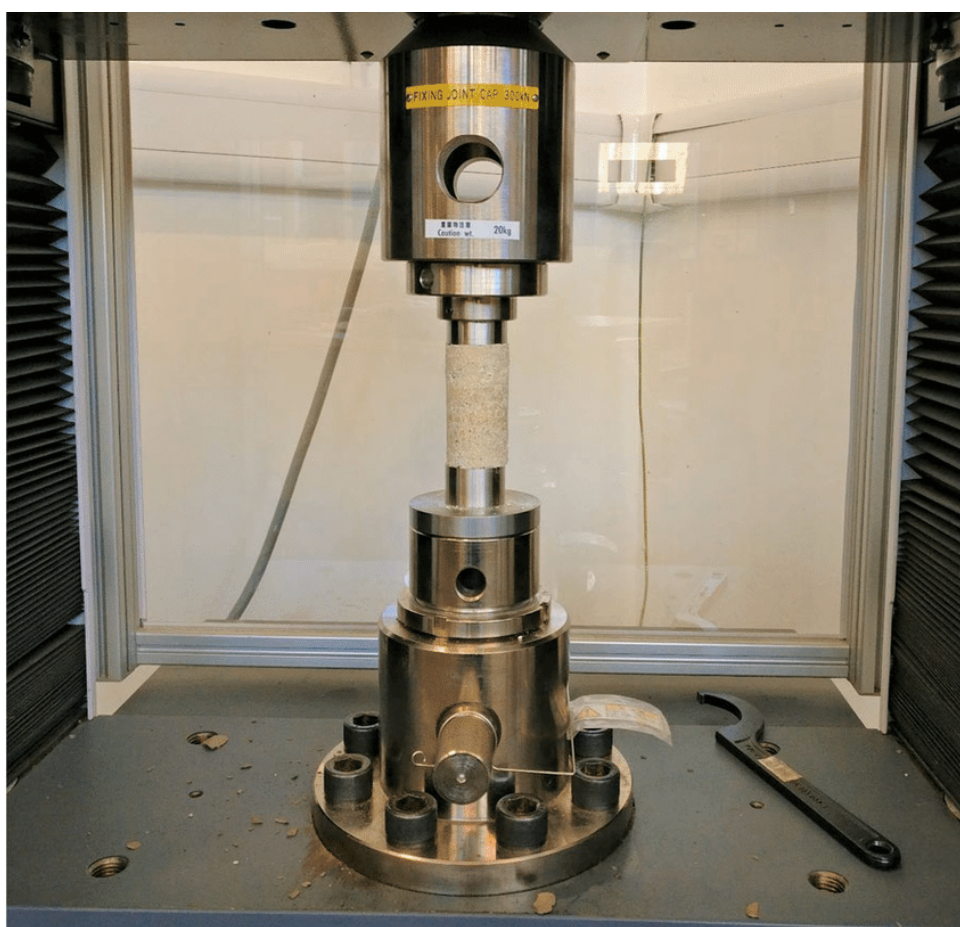
O ensaio de compressão simples, também conhecido como ensaio de compressão não confinada, é utilizado para determinar a resistência de solos coesivos. Para solos estabilizados, a amostra deve ser seca ao ar, destorroada e homogeneizada até obter uma massa representativa de solo. A quantidade de água adicionada deve ser de tal forma que,

a amostra obtenha de 0,5 a 1% à mais de umidade do que ao seu teor de umidade ótima. A moldagem do corpo de prova se dá em um molde cilíndrico metálico com dimensões padronizadas, e compactado com soquete metálico conforme normatizado (73).

Anteriormente ao procedimento, os corpos de prova são colocados em câmara úmida para realização da cura, e após imersos em água durante quatro horas. No ensaio aplica-se uma carga axial numa taxa constante de (140 ± 70) kPa/s, a um corpo de prova cilíndrico até que ocorra a ruptura. A razão entre a carga máxima e a área inicial da seção transversal fornece a tensão de ruptura, correspondente à resistência à compressão simples do solo. Esse ensaio apresenta características específicas, pois o corpo de prova é submetido a uma condição não adensada e não drenada. A aplicação da carga ocorre em uma taxa relativamente rápida, com duração total geralmente entre dez e quinze minutos. Como não há drenagem durante o ensaio, as pressões neutras geradas pelo carregamento não se dissipam, refletindo as condições de campo de curto prazo em solos saturados (70, 67, 74).

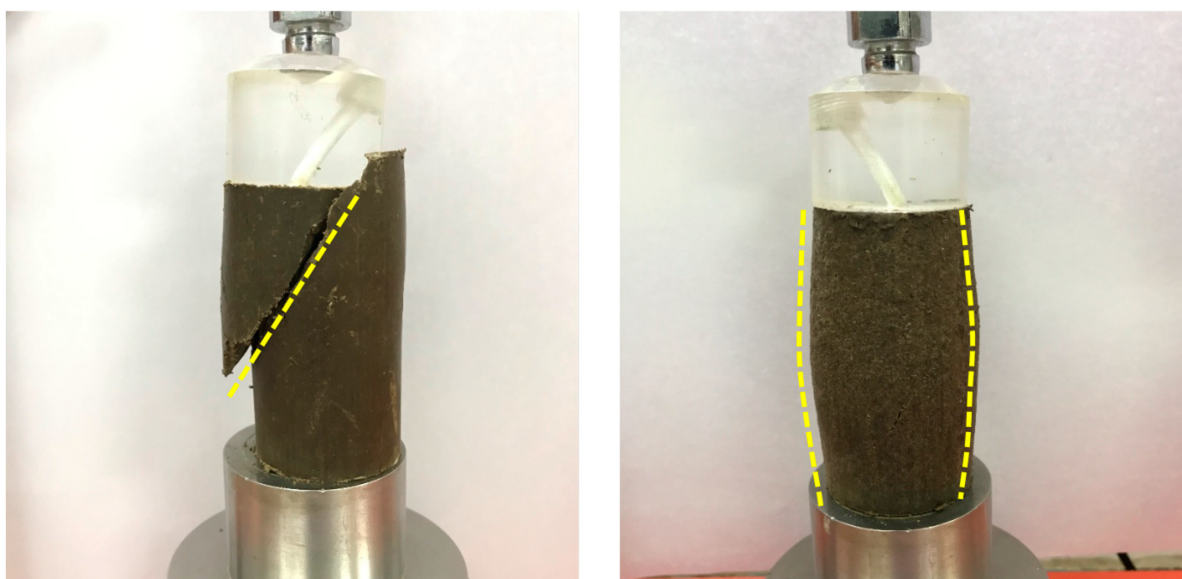
A Figura 2 ilustra um equipamento utilizado no ensaio de compressão simples, e a Figura 3 mostra corpos de prova de argila após o ensaio.

Figura 2 – Equipamento para realização do ensaio de compressão simples.



Fonte: Muguda et al. (2022).

Figura 3 – Ensaio de compressão simples em corpos de prova de argila, à esquerda argila pura, e à direita argila estabilizada com biopolímero.



Fonte: Ying et al. (2025).

O ensaio de compressão simples é particularmente adequado para a análise de solos estabilizados, pois permite estimar a resistência sob baixos níveis de confinamento. Esse ensaio fornece informações sobre o grau de coesão e sobre a contribuição das ligações entre partículas para a resistência mecânica do solo, sendo amplamente utilizado na caracterização de solos estabilizados por adição de cal, cimento e materiais pozolânicos (77, 67).

Vale destacar que os resultados obtidos em ensaios de compressão simples em solos estabilizados são influenciados por diversos fatores. Entre eles, destacam-se o tipo do solo; o tipo e a quantidade de estabilizante; o estado de adensamento do solo, representado pelo índice de vazios; a sensibilidade da estrutura do material, associada ao índice de plasticidade; a densidade relativa; as condições de drenagem durante o ensaio; a velocidade de aplicação das cargas; o desenvolvimento de pressões neutras; e o teor de umidade das amostras. Esses fatores podem alterar de forma significativa a resistência medida e devem ser considerados na interpretação dos resultados (67, 77).

2.1.2 ESTABILIZAÇÃO QUÍMICA

A estabilização de solos naturais tem sido estudada como uma alternativa de menor custo e com benefícios ambientais para aplicação em obras de engenharia. O objetivo desse processo é de modificar propriedades físicas e mecânicas do solo para atender às exigências de resistência, rigidez e durabilidade em aplicações geotécnicas. Diversas técnicas são descritas na literatura, incluindo métodos físicos, mecânicos, hidráulicos, térmicos

e químicos, que visam melhorar o desempenho do solo como material de construção. A estabilização química é uma das técnicas mais utilizadas e baseia-se na adição de materiais ligantes ou reagentes ao solo. Entre os aditivos mais comuns estão a cal e o cimento, além de materiais pozolânicos naturais ou industriais. Esses materiais reagem com os minerais do solo, promovendo processos de cimentação e alteração da estrutura interna, o que resulta em aumento da resistência e redução da deformabilidade. Esse tipo de estabilização é amplamente empregado em obras de engenharia civil, como muros de contenção, fundações e outras estruturas de terra (77, 3, 4).

A estabilização por mistura solo-cal e solo-cimento é amplamente empregada em camadas de fundação, estruturas de contenção e obras de terra. A cal atua, inicialmente, por troca iônica e floculação das partículas finas, seguida de reações pozolânicas de médio e longo prazo. O cimento atua como ligante hidráulico, formando produtos cimentantes que unem as partículas do solo e aumentam a resistência mecânica desde os primeiros períodos de cura (77).

Contudo, a produção contínua de cimento Portland acarreta impactos ambientais significativos, em especial devido à emissão de dióxido de carbono e ao elevado consumo energético durante sua fabricação (5). Esses fatores têm incentivado a busca por alternativas mais sustentáveis, com destaque para o uso de resíduos industriais e agrícolas como substitutos parciais desses materiais.

A cinza da casca de arroz, subproduto de atividades agrícolas, tem sido estudada como alternativa para uso na estabilização de solos. Esse material apresenta elevado teor de sílica amorfa e comportamento pozolânico, podendo ser utilizado como aditivo ou como substituto parcial do cimento em misturas solo-cimento. O uso desse resíduo permite a melhoria das propriedades mecânicas do solo estabilizado e contribui para a redução do consumo de ligantes convencionais (6, 2). Além disso, a cinza da casca de arroz apresenta propriedades pozolânicas mais expressivas quando incinerada sob condições controladas, sendo recomendável a utilização de temperaturas em torno de 500 °C por períodos próximos de 120 minutos, de modo a favorecer a obtenção de elevada finura e teor de sílica (52, 78).

O comportamento dos solos sedimentares está associado às forças transmitidas nos contatos entre partículas e ao atrito mobilizado durante o deslocamento relativo. A reorientação e distorção das partículas controlam as propriedades de deformabilidade e resistência em solos coesivos. Em solos estabilizados quimicamente há a presença de ligações coesivas nos contatos entre partículas. Nessas condições, os deslocamentos iniciais são resistidos por essas ligações antes da mobilização do atrito entre partículas (77, 67). Essa parcela adicional de resistência é associada à coesão decorrente da ligação entre solo e aditivo, e difere da coesão aparente observada em argilas parcialmente saturadas associada à pressão neutra (poro-pressão da água) negativa (67, 68).

2.1.3 ESTABILIZAÇÃO POR COMPACTAÇÃO E POR REFORÇO

A técnica de estabilização de solos por compactação profunda fundamenta-se no incremento da densidade de depósitos de solos sem coesão para mitigar a ocorrência de recalques excessivos e prevenir a liquefação sob carregamentos dinâmicos. Diferente dos métodos de compactação superficial, esta abordagem utiliza processos dinâmicos para desestruturar o arranjo inicial das partículas, permitindo que estas se reorganizem em configurações mais estáveis e compactas. O mecanismo de densificação varia conforme o grau de saturação, no qual em materiais saturados a melhoria é alcançada pela indução da liquefação e subsequente dissipação da pressão neutra; enquanto em solos parcialmente saturados o processo ocorre predominantemente pelo colapso da estrutura e expulsão do ar dos vazios, muito similar à compactação Proctor feita em laboratório (77, 79).

A técnica de estabilização de solos por reforço fundamenta-se na inserção de elementos resistentes à tração ou compressão, integrados ao maciço durante a execução de aterros ou instalados em campo para o fortalecimento do terreno sob estruturas existentes. Diferentemente de abordagens que visam a modificação das propriedades inerentes do solo, como a estabilização química, o reforço concentra-se no aprimoramento do desempenho do sistema solo-aditivo. O objetivo principal desta metodologia é o incremento da resistência ao cisalhamento aparente e a mitigação de deformações internas e de contorno, permitindo a viabilização de obras em locais com condições geotécnicas desfavoráveis para suporte de fundações e estruturas de terra (77, 79).

Dentre as metodologias aplicadas, destaca-se a terra armada, caracterizada pela utilização de armaduras horizontais para suportar exclusivamente esforços de tração em estruturas erguidas por camadas sucessivas. Já a técnica de solo grampeado emprega barras de aço injetadas com calda de cimento para a estabilização de taludes e escavações profundas, enquanto as estacas raiz formam uma massa reforçada para suporte estrutural e contenção. Complementarmente, o uso de colunas de pedra proporciona suporte vertical e aceleração da drenagem em solos moles, e os geotêxteis atuam como inclusões extensíveis que elevam o confinamento e a rigidez do sistema, sendo amplamente aplicados em aterros sobre solos compressíveis e na pavimentação de vias temporárias (77, 79).

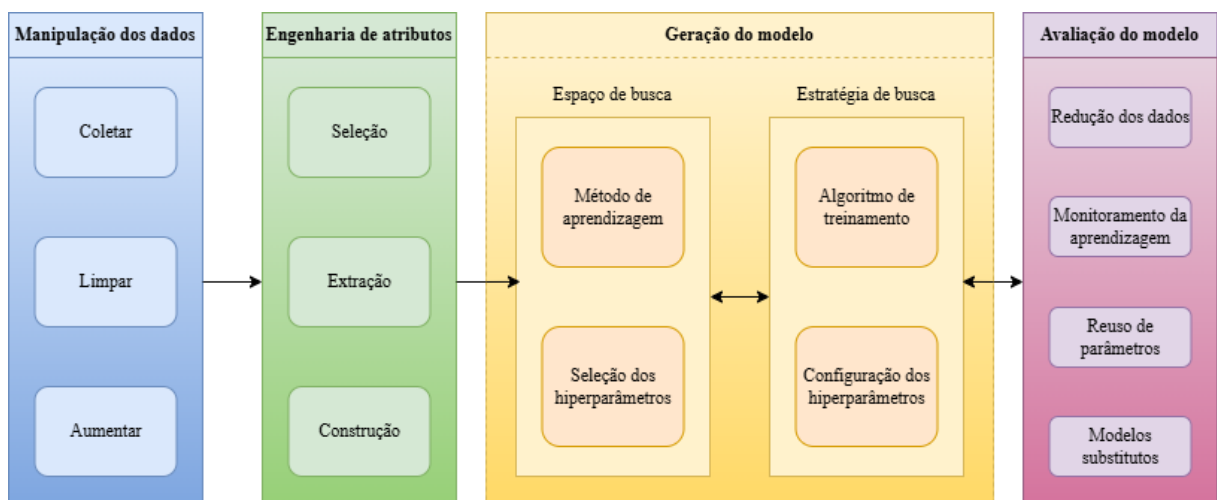
2.2 APRENDIZADO DE MÁQUINA

O aprendizado de máquina consiste no desenvolvimento e uso de modelos capazes de prever novos resultados a partir de informações previamente observadas, ou seja, são modelos que “aprendem” com os dados. De modo geral, esses modelos podem ser classificados em duas categorias principais. No aprendizado supervisionado, os dados possuem rótulos conhecidos, e o algoritmo detecta padrões entre a relação das entradas e das respostas fornecidas. Já no aprendizado não supervisionado, não há categorização prévia, e o algoritmo busca identificar padrões e estruturas internas do conjunto de

dados com base nos atributos disponíveis (80, 81). Atributos (ou *features*) podem ser compreendidos como variáveis de entrada do modelo ou, de forma mais precisa, como “qualquer informação que fornecemos ao nosso modelo” (Grus, 2015, p. 148).

O processo de aprendizado de máquina pode ser dividido em algumas etapas essenciais (Figura 4). A primeira etapa envolve a análise inicial dos dados, que ajuda a entender sua estrutura e possíveis padrões. Em seguida, ocorre o tratamento dos dados, no qual o conjunto é preparado para uso nos modelos por meio de limpeza, remoção de valores fora do esperado, aumento de dados e criação ou transformação de atributos. Após, vem a fase de construção dos modelos, que inclui a escolha dos algoritmos e o ajuste de seus hiperparâmetros. Por fim, os resultados podem apoiar diferentes tipos de decisões. Como descrito na literatura, o algoritmo de busca procura boas configurações de aprendizado dentro do espaço disponível a partir da avaliação dessas configurações. Por esse motivo, há uma relação de iteração entre as etapas de geração e avaliação do modelo, ilustrada pelas setas bidirecionais no fluxograma (56, 82, 83).

Figura 4 – Fluxograma das etapas do aprendizado de máquina.



Fonte: Adaptado de He et al. (2021), Baratchi et al. (2024) e Shen et al. (2024).

2.2.1 MANIPULAÇÃO DOS DADOS

A etapa de manipulação ou preparação dos dados pode ser dividida em três etapas: coleta, limpeza e aumento dos dados.

Há um consenso de que a qualidade dos conjuntos de dados possui forte impacto no desempenho de modelos de aprendizado de máquina. Apesar disso, localizar conjuntos de dados com boa qualidade é uma tarefa difícil. A internet permite pesquisar uma grande variedade de bases porém, esse processo pode apresentar limitações, como o uso de palavras-chave que não retornam resultados adequados e conjuntos de dados que não estão categorizados de forma correta (56).

Para contornar essas dificuldades, ou em situações em que não é possível obter novos dados de forma direta (como ocorre em testes de direção autônoma), a simulação de dados surge como uma alternativa prática. Nesse caso, criam-se dados sintéticos que procuram reproduzir, da melhor forma possível, o comportamento observado no mundo real (56).

Mesmo assim, métodos convencionais de coleta continuam sendo amplamente usados, principalmente em áreas experimentais. Em geotecnia, a obtenção de amostras físicas é um procedimento comum, como relatado nos estudos de Ngo et al. (49), Priyadarshee et al. (50), Mozumder and Laskar (51), Tabarsa et al. (52), Mahmoodzadeh et al. (53) e Zhang et al. (54). Esses conjuntos de dados também são utilizados no presente trabalho, conforme detalhado na Seção 3.

Uma das primeiras tarefas após a obtenção dos dados é verificar como eles estão organizados, para que possam ser lidos corretamente em um ambiente de análise. Em muitos casos, também é preciso reunir informações que vêm de fontes distintas e com formatos diferentes. Quando os dados possuem dimensões ou escalas de tempo diferentes, é necessário ajustá-los para manter a coerência. Há situações em que o dado bruto não segue um formato estruturado (dados não estruturados). Nesses casos, é preciso extrair as informações relevantes e organizá-las em tabelas antes de seguir para as etapas seguintes (84).

O processo de limpeza de dados tem como objetivo localizar e tratar casos de valores incompletos, inconsistentes, imprecisos ou duplicados, de modo a produzir conjuntos de dados com melhor qualidade. Uma tarefa frequente na limpeza de dados é a padronização, cujo propósito é transformar registros que aparecem de formas diferentes em um formato único e coerente. Isso reduz ambiguidade e facilita as etapas seguintes do processamento (84, 85).

Valores incompletos correspondem aos dados ausentes no conjunto. Valores inconsistentes aparecem quando um registro não condiz com a categoria à qual pertence, ou quando há repetições indevidas de informações que surgem em arquivos diferentes e em vários formatos. Valores imprecisos estão ligados a erros que impedem a identificação correta do dado e prejudicam sua confiabilidade. Por fim, valores duplicados são registros inseridos mais de uma vez por engano (84, 85).

A etapa de aumento de dados é útil quando o conjunto disponível não representa bem o comportamento real do problema. Isso pode ocorrer quando certas faixas de valores aparecem pouco ou nem chegam a ser observadas, criando um desbalanceamento que afeta o ajuste dos modelos. Situações assim podem surgir por ruído na coleta, por limitações do método de amostragem ou por diferenças entre a amostra e a população real. Quando os valores pouco observados são importantes para o processo de predição, esse desbalanceamento tende a prejudicar ainda mais os resultados, podendo causar *overfitting*

no modelo de treinamento (86, 87, 56).

Para reduzir esses efeitos, o aumento de dados busca gerar novos exemplos que ampliem o conjunto original sem alterar o significado das classes ou variáveis de interesse. Essa prática é comum em áreas como classificação de imagens, nas quais se cria versões modificadas de uma mesma amostra com o intuito de ampliar a diversidade e reduzir problemas de sobreajuste. No entanto, é preciso atenção para que as alterações feitas não descaracterizem o dado original. Ajustes inadequados produzem amostras que não representam a categoria correta, dessa forma, o aumento de dados exige escolhas cuidadosas para equilibrar diversidade e fidelidade (88).

2.2.2 ENGENHARIA DE ATRIBUTOS

A etapa de engenharia de atributos pode ser dividida em três etapas: seleção, extração e construção de atributos.

A seleção de atributos busca definir um subconjunto capaz de simplificar o modelo, reduzir o risco de *overfitting* e melhorar o desempenho das previsões. Em geral, procura-se escolher variáveis pouco correlacionadas entre si, mas com relação forte com a variável alvo. Os métodos de busca são divididos em três grupos: busca completa, busca heurística e busca aleatória. A avaliação dos atributos pode ser feita por meio de um alvo definido e de um critério de pontuação, como variância, coeficiente de correlação e teste qui-quadrado. Esse processo de escolha e avaliação ocorre de forma iterativa até que se obtenha um subconjunto consistente de atributos (89, 56).

A extração e construção de atributos tem os mesmos objetivos da seleção, porém, são processos que modificam as variáveis originais. A extração de atributos reduz a dimensionalidade dos atributos através de funções de mapeamento, nas quais as técnicas mais usadas são a *Principal Component Analysis* (PCA), *Independent Component Analysis* (ICA), *isomap*, métodos de redução não linear de dimensionalidade e a *Linear Discriminant Analysis* (LDA). A construção de atributos cria novas variáveis a partir das originais com o objetivo de melhorar a capacidade de generalização do modelo (e por consequência seu desempenho), e entre os métodos comuns estão transformações de pré-processamento, como padronização, normalização e discretização (56).

2.2.3 GERAÇÃO DO MODELO

A etapa de geração do modelo pode ser dividida em duas etapas: espaço e estratégia de busca.

A etapa de geração de modelos em AutoML inicia com a organização das decisões que compõem uma solução de aprendizado de máquina. Para isso, é formado um espaço de busca que reúne todas as escolhas possíveis, como formas de tratar os dados, métodos de aprendizagem e ajustes desses métodos. Esse espaço precisa ser amplo o bastante

para que o processo encontre soluções com bom desempenho, mas não pode ser grande demais a ponto de tornar a busca lenta. O ponto inicial é a separação dos elementos que formam uma solução comum de aprendizado de máquina em pequenas partes. A partir disso, esse conjunto pode ser reorganizado em uma versão mais compacta, ou até mesmo transformado por funções que tornem a busca mais simples. Esses passos adicionais não são obrigatórios, mas ajudam quando o espaço de busca é muito grande ou difícil de explorar (83).

O espaço de busca inclui a escolha do algoritmo de aprendizagem e os hiperparâmetros que acompanham cada opção. A seleção do algoritmo é tratada como uma escolha entre categorias, e cada categoria ativa apenas os hiperparâmetros que são relevantes para cada método. Isso cria uma estrutura em formato de árvore, na qual certas opções só aparecem quando outras são escolhidas. Alguns hiperparâmetros também influenciam o desempenho de forma diferente ao longo de seus valores, ou seja, pequenas mudanças em taxas de aprendizagem muito baixas podem afetar o modelo mais do que variações em valores muito altos. Dessa forma, em muitos casos, é útil transformar o intervalo desses hiperparâmetros antes da busca, como no uso de escalas logarítmicas. Assim, a seleção de valores passa a considerar melhor essas diferenças, o que torna a busca por configurações adequadas mais equilibrada (82).

Após a definição do espaço de busca, é necessário estabelecer as configurações que orientam como o algoritmo procura as melhores soluções, ou seja, decide-se como a pesquisa será conduzida. O método escolhido deve apresentar bom desempenho, mas também precisa ser viável do ponto de vista computacional, evitando custos excessivos durante o processo (83, 82). Estudos mostram que é possível, a partir de uma modelagem geral de ML, obter melhores desempenhos dos modelos através de uma específica escolha dos hiperparâmetros, ou seja, modelos com conjuntos de dados iguais podem obter melhores resultados, à depender das configurações dos hiperparâmetros (90, 91).

2.2.4 AVALIAÇÃO DO MODELO

A etapa de avaliação do modelo tem como finalidade selecionar, de forma iterativa, os melhores candidatos obtidos pelo método de busca. Os modelos escolhidos podem ser treinados até atingirem a convergência, permitindo verificar o desempenho de cada um e identificar aquele que apresenta o melhor resultado para a tarefa estudada. Como esse processo geralmente demanda alto custo computacional, recorrem-se a técnicas que aceleram o treinamento sem prejudicar a qualidade final das previsões. A avaliação do modelo pode ser dividida em quatro categorias: redução de dados, monitoramento das curvas de aprendizagem, reuso de parâmetros, e utilização de modelos substitutos. (83, 56, 82).

A técnica de redução de dados pode ser aplicada de diferentes formas, mas todas

têm o mesmo objetivo: reduzir a quantidade de dados utilizada pelo modelo na fase de treinamento. Esse objetivo pode ser alcançado por vários meios, como o uso de uma subamostra do conjunto de dados no treinamento; a redução do tempo de treinamento; a limitação do processo de treinamento, com a redução do número de iterações; a diminuição da qualidade dos dados, como em tarefas de visão computacional; e a redução da dimensão do conjunto de dados (56, 82).

Resultados apresentados por Zela et al. (92) indicam que, o desempenho obtido após períodos curtos de treinamento apresenta pouca relação com aquele observado após treinamentos mais longos. Esse resultado sugere que, em alguns casos, buscas extensas por configurações de redes neurais não são necessárias para obter modelos com desempenho satisfatório.

Ao acompanhar o processo de treinamento, é possível avaliar o desempenho do modelo e definir critérios de parada com base nos estágios iniciais, ou ainda estimar o desempenho a partir das informações disponíveis no momento. O objetivo é, além de evitar *overfitting*, encerrar o treinamento quando não houver ganho relevante de desempenho em um determinado intervalo de tempo (*early stopping*) ou quando a estimativa atual de desempenho não for satisfatória (*learning curve exploration*) (93, 94, 81, 83).

A reutilização de parâmetros consiste no uso de parâmetros que apresentam melhor desempenho nos estágios iniciais do treinamento. Realiza-se uma busca inicial por conjuntos de parâmetros promissores, e o modelo resultante (pré-treinado) é ajustado de forma incremental para acelerar o processo de avaliação. Em redes neurais, esses parâmetros podem ser interpretados como os pesos da rede (95, 96, 97).

Modelos substitutos (menores ou mais simples) podem ser utilizados para tornar o processo de avaliação mais eficiente, por meio da estimativa do desempenho do modelo sob diferentes configurações dos parâmetros do algoritmo de aprendizado. Entretanto, é necessário verificar a relação entre os parâmetros de aprendizado e o desempenho obtido, para evitar prejuízos nos resultados do treinamento e, conseqüentemente, da validação do modelo (83, 56).

2.3 SISTEMA DE APRENDIZADO DE MÁQUINA AUTOMATIZADO

Nas últimas décadas, técnicas de aprendizado de máquina (incluindo abordagens mais sofisticadas, como modelos de aprendizado profundo) tornaram-se amplamente utilizadas em diversos domínios. Concomitantemente, cresceu o interesse em automatizar essas etapas, o que consolidou o campo conhecido como AutoML. Essa área representa, atualmente, uma das aplicações mais consolidadas e impactantes da automação dentro do processo de ciência de dados (84, 57).

A Figura 5 organiza os principais elementos envolvidos em problemas de AutoML

em três grandes grupos. A partir dessa estrutura, destacam-se técnicas frequentemente aplicadas tanto na etapa de geração quanto na fase de avaliação dos modelos.

Figura 5 – Resumo de grupos de problemas em AutoML, com exemplos de espaço de busca, estratégias de busca e formas de avaliação.

Problema em AutoML	Espaço de busca	Estratégia de busca	Avaliação
Configuração de <i>pipelines</i>	Conjuntos de hiperparâmetros usados em <i>pipelines</i> , como taxa de aprendizado e tamanho do lote	Exemplos incluem busca aleatória e métodos baseados em probabilidade	Parada antecipada e análise de curvas de aprendizado
Projeto de arquiteturas de redes	Escolhas de operações e possíveis ligações entre blocos da rede	Aprendizagem por reforço e descida de gradiente	Compartilhamento de pesos e modelos que estimam o desempenho
Modelos base (<i>foundation models</i>)	Conjuntos de métodos de pré-treino e ajustes finos, além de parâmetros e hiperparâmetros	Métodos baseados em gradiente e busca probabilística	Preditores de desempenho e métodos <i>bandit-based</i>

Fonte: Adaptado de Shen et al. (2024).

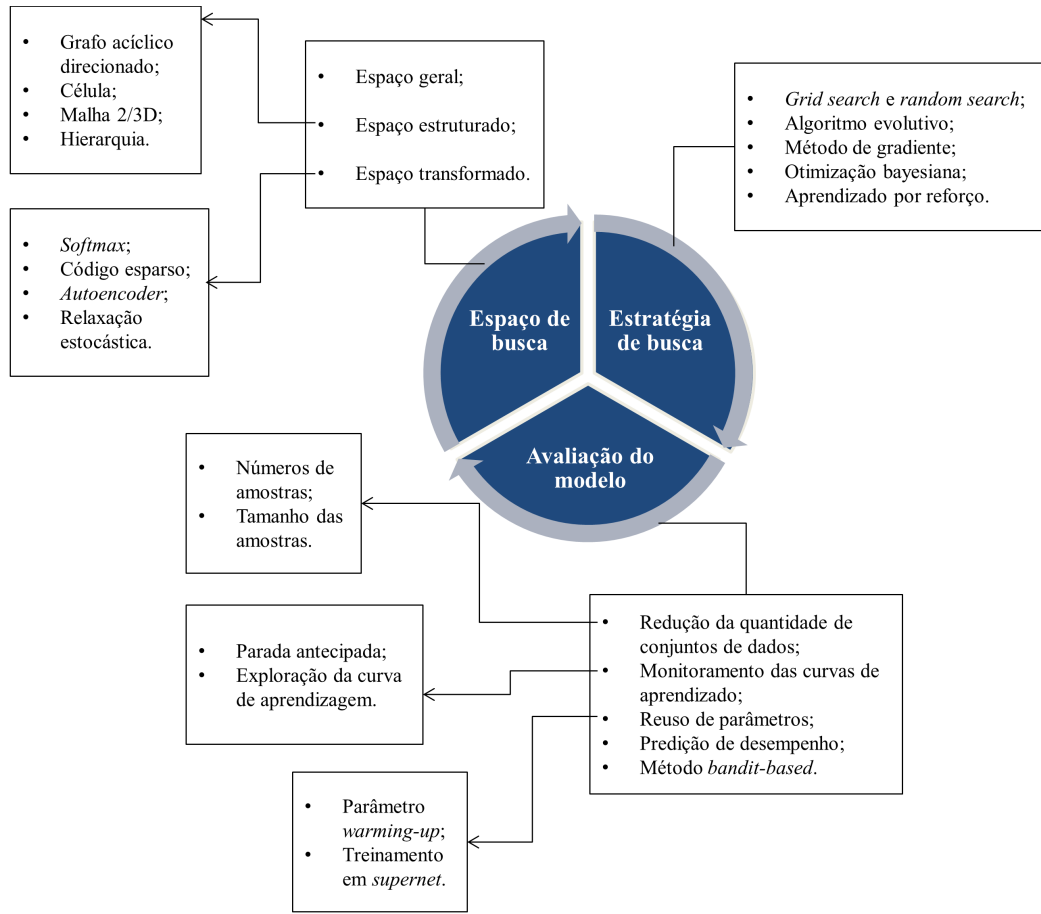
A Figura 6 mostra um detalhamento das atuais abordagens realizadas em AutoML para as etapas de geração e avaliação dos modelos.

Um panorama comparativo dos sistemas de AutoML analisados é apresentado na Tabela 1, enquanto as seções seguintes discutem com mais detalhes seus atributos de desempenho e suas capacidades funcionais.

2.3.1 AUTOGLUON

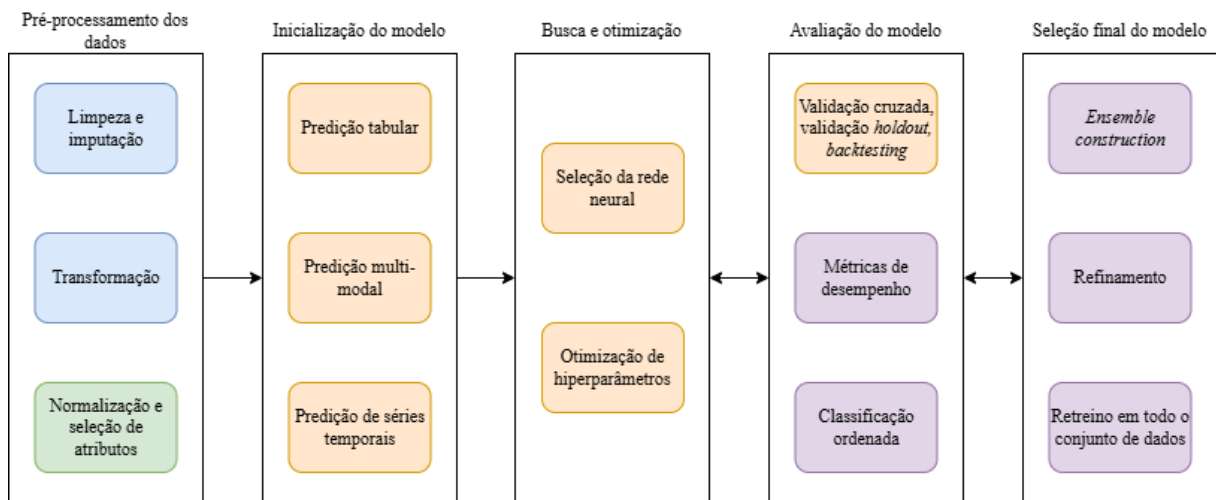
O AutoGluon para dados tabulares, é uma ferramenta escrita em Python cujo foco é a automação do processo de treinamento de modelos. A estrutura do sistema permite que diferentes modelos sejam ajustados em camadas sucessivas, formando conjuntos que buscam bom desempenho dentro do tempo definido pelo usuário. O sistema identifica o tipo de dado presente em cada coluna e aplica tratamentos adequados para cada formato, o que reduz falhas comuns na etapa de preparação (98). A Figura 7 apresenta um esquema geral do funcionamento do AutoGluon.

Figura 6 – Denominações das abordagens de AutoML para a geração e avaliação de modelos.



Fonte: Adaptado de Shen et al. (2024).

Figura 7 – Fluxograma esquemático do AutoGluon.



Fonte: Elaborado pelo autor (2026).

Tabela 1 – Comparativo dos sistemas de AutoML.

Sistema	Ano de lançamento	Técnica de otimização	Ferramenta de ML	Meta-aprendizagem	Pós-processamento
AutoGluon (98)	2020	<i>Ensemble-Based Search</i>	MXNet, PyTorch	Sim	Sim
AutoKeras (99)	2019	<i>Bayesian Optimization</i>	Keras, TensorFlow	Não	Sim
FLAML (100)	2021	<i>Cost-Aware Search</i>	flaml	Não	Não
H2O (101)	2020	<i>Random and Ensemble-Based</i>	h2o	Sim	Sim
TPOT (102)	2016	<i>Genetic Programming</i>	scikit-learn	Não	Sim

Fonte: Elaborado pelo autor (2025).

2.3.2 AUTOKERAS

O AutoKeras foi criado pelo DATA Lab da Texas A&M University. Ele automatiza a construção e o ajuste de modelos baseados em redes neurais. Um ponto positivo é que o sistema pode ser executado em computadores pessoais sem dependência de contêineres ou serviços externos. Assim, o processo ocorre de forma simples, sem configurações complexas. O sistema também reduz preocupações relacionadas à privacidade, pois o fluxo inteiro pode ser executado localmente. Por ser gratuito, o Auto-Keras se torna uma alternativa acessível para estudos e experimentações acadêmicas (99).

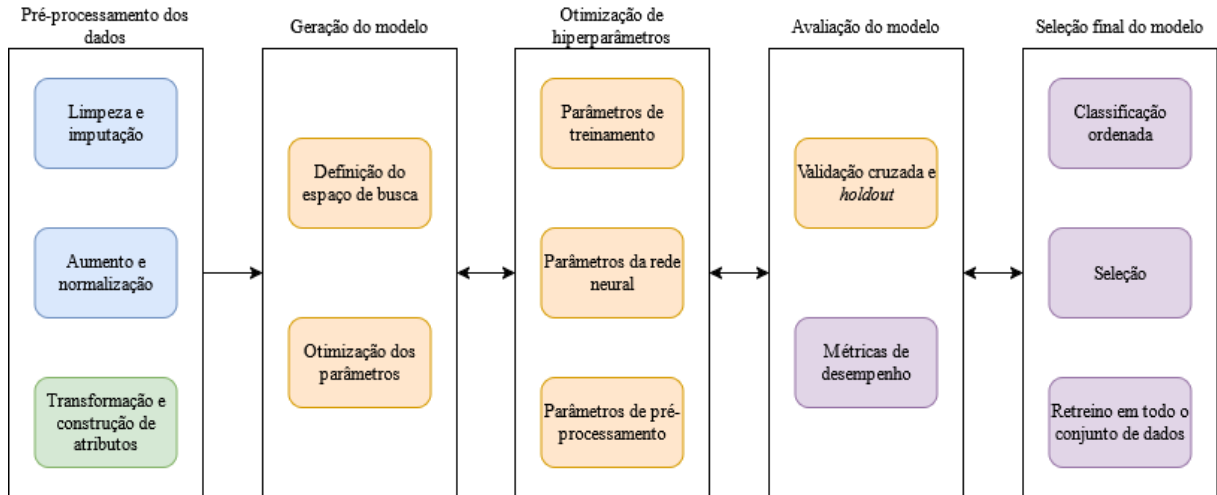
A Figura 8 apresenta a estrutura do AutoKeras. O sistema ajusta o uso de CPU, GPU e memória ao manter apenas os dados ativos na RAM. Para cada tarefa, a API acionada pelo usuário troca informações com módulos internos que controlam o fluxo do processo. A busca por configurações é conduzida por otimização Bayesiana, que utiliza processos Gaussianos e rotinas de ajuste baseadas em CPU.

2.3.3 FAST LIGHTWEIGHT AUTOML (FLAML)

O FLAML é uma biblioteca em Python cujo propósito é executar AutoML com baixo custo computacional. Em vez de percorrer todo o espaço possível de configurações, o sistema concentra a busca em combinações que tendem a apresentar bom desempenho com tempo reduzido. Durante o processo, o FLAML escolhe o algoritmo, ajusta hiperparâmetros, define o tamanho das amostras e o método de validação, considerando o impacto direto

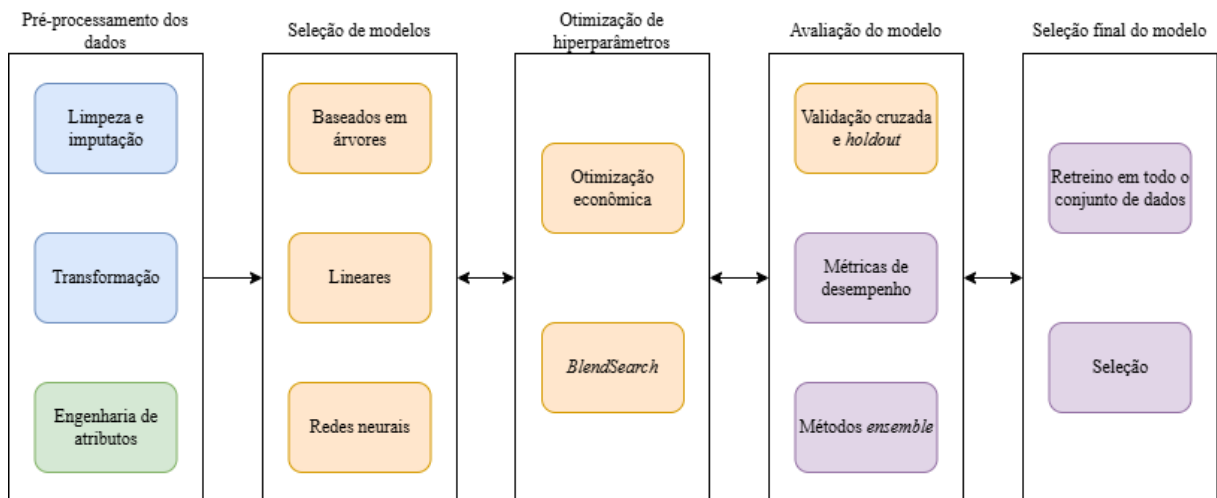
no tempo de execução e na qualidade das predições (100). A Figura 9 mostra uma visão geral da estrutura da ferramenta.

Figura 8 – Fluxograma esquemático do AutoKeras.



Fonte: Elaborado pelo autor (2026).

Figura 9 – Fluxograma esquemático do FLAML.



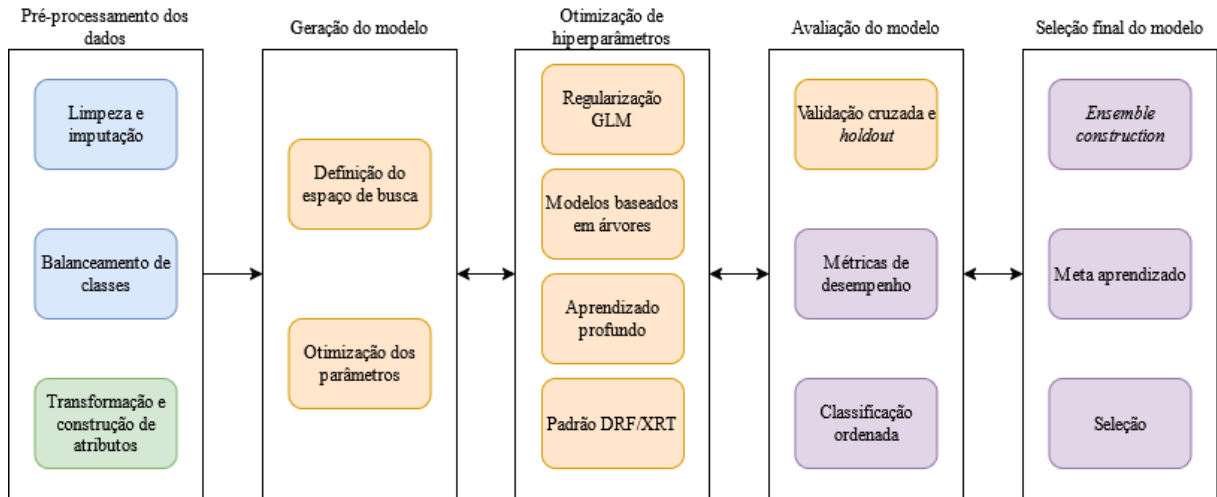
Fonte: Elaborado pelo autor (2026).

2.3.4 H2O

O H2O AutoML integra o ecossistema H2O e oferece recursos para aprendizado supervisionado em problemas de regressão e classificação. O sistema aplica etapas básicas de tratamento dos dados, como imputação de valores ausentes, normalização quando necessária e codificação de variáveis categóricas. Modelos baseados em árvores conseguem lidar diretamente com categorias, o que reduz transformações adicionais. A interface do

H2O AutoML solicita apenas o conjunto de dados e a variável alvo, permitindo também definir limites de tempo ou de quantidade de modelos (101). A Figura 10 resume a organização interna da ferramenta.

Figura 10 – Fluxograma esquemático do H2O.

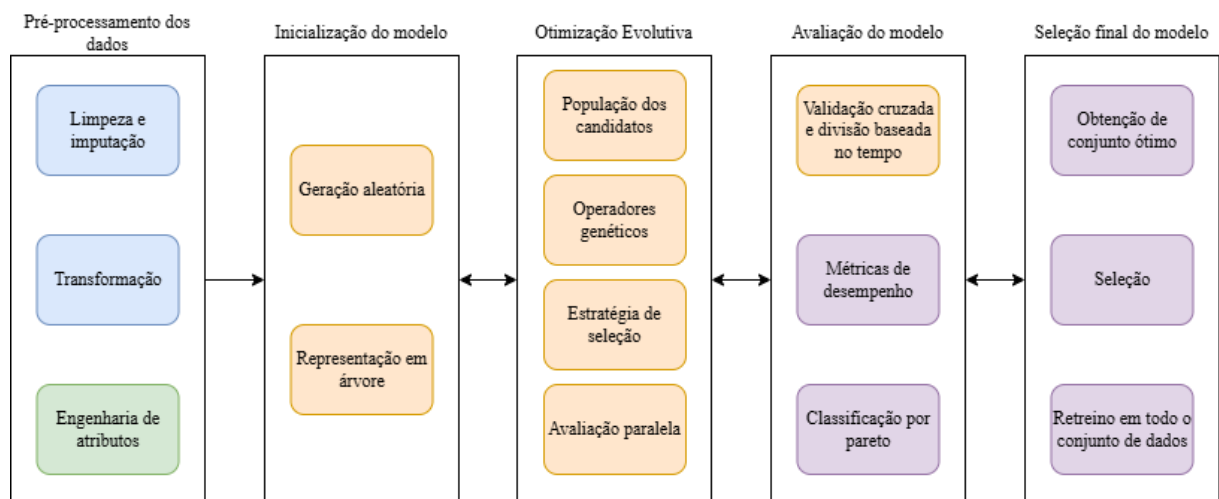


Fonte: Elaborado pelo autor (2026).

2.3.5 TREE-BASED PIPELINE OPTIMIZATION TOOL (TPOT)

O TPOT utiliza GA para construir e ajustar *pipelines* de aprendizado de máquina. No início da execução, o sistema cria uma população inicial de *pipelines*. A cada etapa, os *pipelines* com maior desempenho são mantidos e usados para formar novas combinações. Esse ciclo segue até atingir o número de iterações configurado pelo usuário ou até ocorrer estabilização dos resultados. No final, o TPOT retorna o *pipeline* com melhor desempenho (102). A Figura 11 apresenta um esquema simplificado desse processo.

Figura 11 – Fluxograma esquemático do TPOT.



Fonte: Elaborado pelo autor (2026).

3 MATERIAIS E MÉTODOS

Este capítulo será dividido em 5 seções: primeiro detalham-se todos os conjuntos de dados utilizados para o treinamento dos modelos; após, descreve-se os principais parâmetros e métodos utilizados pelos sistemas de AutoML; em seguida, informam-se as métricas de desempenho utilizadas para avaliar os resultados; após, explica-se a análise de importância de atributos utilizada; e por fim, descrevem-se as configurações computacionais para a execução dos testes.

3.1 CONJUNTOS DE DADOS

Tabela 2 – Resumo dos conjuntos de dados.

Conjunto de dados e Referência	Nº de amostras	Nº de variáveis	Tipo de solo*/rocha	Estabilizante	País
D1 (48)	179	7	Diversos ¹	Cal	EUA
D2 (48)	60	7	Diversos ²	Cimento	EUA
D3 (49)	100	14	Diversos ³	Cimento	Vietnã
D4 (50)	126	5	CL	Diversos ⁴	Índia ⁵
D5 (51)	283	8	CH e CL	Geopolímero	Índia
D6 (3)	189	9	Diversos ⁶	Diversos ⁷	Diversos ⁸
D7 (52)	137	6	SM e MH	CLR ⁹	Malásia
D8 (53)	170	4	Diversos ¹⁰	– ¹¹	Irã
D9 (20)	351	7	Areias	MICP ¹²	Diversos ¹³
D10 (54)	150	3	CH	Diversos ¹⁴	China

* Solos classificados de acordo com o USCS são indicados por símbolos; nesses casos, consultar a Figura 12; ¹ CH, OL, CL, ML, OH, SC, SM, GC e GM; ² SP-SM, CH, SM, ML, CL, GC, GM e SC; ³ argila amarela, argila arenosa orgânica preta e argila arenosa preta; ⁴ cinza de lagoa, cinza de casca de arroz e cimento; ⁵ considerando que a Índia possui numerosos depósitos de argila caulínica, um mineral amplamente extraído no país, e que os autores são afiliados a instituições indianas, é razoável inferir que os experimentos e a coleta de amostras foram provavelmente realizados na Índia; ⁶ CL, ML, MH, CH e CL-ML; ⁷ cal, cimento, materiais pozolânicos e cinza volante; ⁸ Argélia, EUA, Turquia, Índia, Bangladesh e Austrália; ⁹ diversas combinações de misturas de cimento, cal e cinza de casca de arroz (CLR); ¹⁰ argilito, granito, xisto, arenito, travertino, calcário, ardósia, dolomito e marga; ¹¹ não há estabilizante porque o estudo realiza ensaios com rochas; ¹² precipitação de calcita induzida microbianamente (*microbially induced calcite precipitation* - MICP); ¹³ Alemanha, Austrália, EUA e China; ¹⁴ cimento e polímero reforçado com fibra de vidro (*glass fiber-reinforced polymer* - GFRP). Fonte: Elaborado pelo autor (2025).

A Tabela 2 apresenta um resumo dos dez conjuntos de dados experimentais utilizados neste estudo. No total, são consideradas 1745 amostras obtidas a partir de ensaios de laboratório, envolvendo diferentes tipos de solos e rochas, materiais estabilizantes, variáveis de entrada adotadas no treinamento dos modelos e condições experimentais oriundas de diversos países. Para os solos, todas as amostras são deformadas e estabilizadas antes da realização dos ensaios. A maioria dos solos foi classificada de acordo com o Sistema

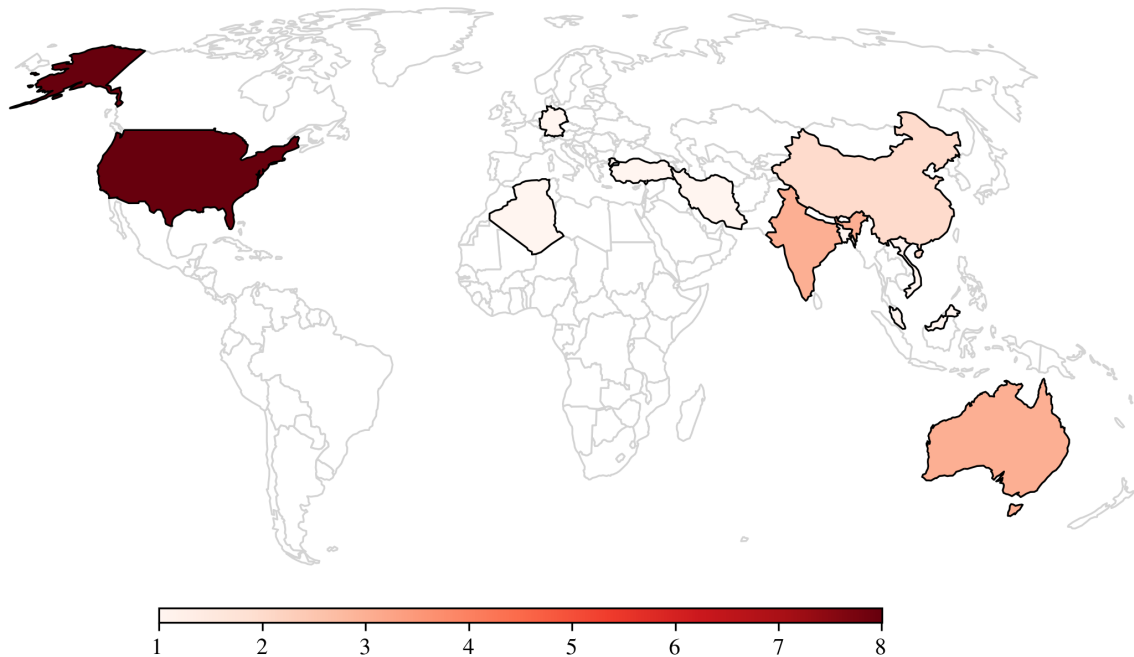
Unificado de Classificação de Solos (*Unified Soil Classification System - USCS*) (103), cujas siglas e descrições estão apresentadas na Figura 12.

Figura 12 – Classificação dos solos de acordo com o USCS.

Sigla do grupo	Nome do grupo
GW	Pedregulhos bem graduados
GP	Pedregulhos mal graduados
GM	Pedregulhos siltosos
GC	Pedregulhos argilosos
SW	Areias bem graduadas
SP	Areias mal graduadas
SM	Areias siltosas
SC	Areias argilosas
ML	Siltos de baixa compressibilidade
CL	Argilas de baixa compressibilidade
OL	Solos orgânicos de baixa compressibilidade
MH	Siltos de alta compressibilidade
CH	Argilas de alta compressibilidade
OH	Solos orgânicos de alta compressibilidade
PT	Solos altamente orgânicos (turfa)

Fonte: Adaptado de ASTM International (2006).

Figura 13 – Distribuição geográfica dos locais das amostras dos conjuntos de dados.



Fonte: Elaborado pelo autor (2025).

Procurou-se conjuntos de dados brasileiros para compor o presente estudo, entretanto não foi encontrado nenhum que atendesse os requisitos de pesquisa, que são:

estudo publicado, resultados com métricas de regressão, dados disponíveis e realização de aprendizado de máquina. O objetivo principal desses requisitos é realizar uma comparação entre os resultados dos estudos de referência com o presente trabalho.

A Figura 13 apresenta a distribuição geográfica das amostras consideradas nos conjuntos de dados. A intensidade da coloração indica a frequência relativa de amostras por país. Tons de vermelho mais escuros representam países com maior número de amostras, enquanto tons mais claros indicam menor quantidade de registros.

As Tabelas 3–7, 9–10 e 12–14, mostram os resultados das principais estatísticas das variáveis dos conjuntos de dados, como média, desvio padrão (DP), percentis (25%, 50% e 75%) e valores máximos (Máx) e mínimos (Mín); e as Figuras 14–23 mostram as distribuições dos valores de cada variável de cada conjunto de dados.

Tabela 3 – Estatísticas básicas do conjunto de dados D1.

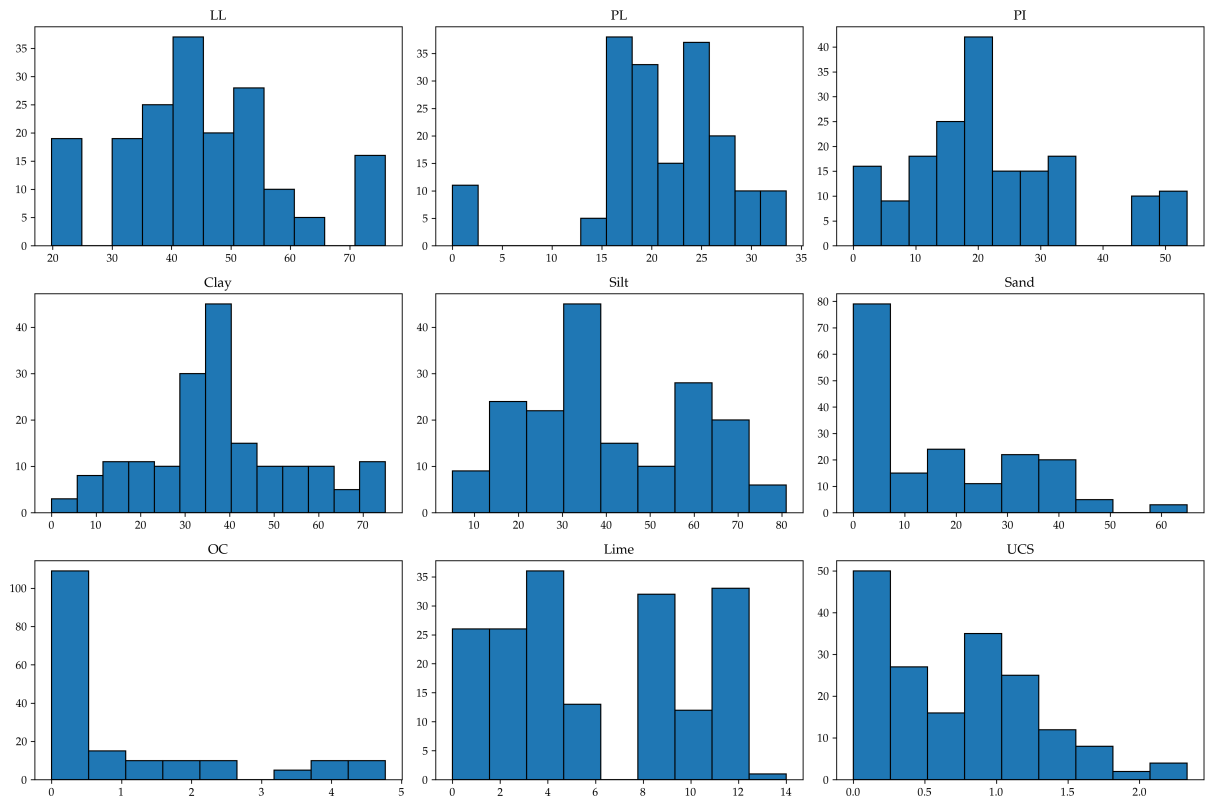
Variável	Nome	Média	DP	Mín	25%	50%	75%	Máx
LL	Limite de liquidez (%)	45,2	13,8	19,8	37,0	43,0	52,1	76,0
PL	Limite de plasticidade (%)	21,2	7,0	0,0	18,0	21,1	25,6	33,5
PI	Índice de plasticidade (%)	22,1	13,0	0,0	14,5	20,5	28,0	53,5
Clay	Teor de argila (%)	37,6	17,2	0,0	29,0	38,5	46,4	75,0
Silt	Teor de silte (%)	41,1	19,3	5,0	30,1	37,0	57,3	81,0
Sand	Teor de areia (%)	16,3	15,9	0,0	1,7	11,7	29,5	65,0
OC	Teor de matéria orgânica (%)	1,0	1,5	0,0	0,0	0,2	1,6	4,8
Lime	Teor de cal (%)	5,9	4,1	0,0	2,0	6,0	10,0	14,0
UCS	Resist. à compressão simples (MPa)	0,7	0,6	0,0	0,2	0,7	1,1	2,3

Fonte: Elaborado pelo autor (2025).

3.1.1 CONJUNTOS DE DADOS D1 E D2

Um dos objetivos do estudo de Gajurel et al. (48) foi consolidar sete conjuntos de dados públicos para o treinamento e a avaliação de algoritmos de aprendizado de máquina, aplicados à predição da RCS de solos estabilizados com cal (conjunto D1) e com cimento (conjunto D2). As variáveis de entrada consideradas incluem os limites de Atterberg – limite de liquidez (LL), limite de plasticidade (PL) e índice de plasticidade (PI) –, parâmetros da distribuição granulométrica – teores de silte (Silt), areia (Sand) e argila (Clay) –, teor de matéria orgânica (OC), além das propriedades e da dosagem do agente estabilizante – teor de cal (Lime) e teor de cimento (Cement).

Figura 14 – Histogramas das variáveis do conjunto de dados D1.



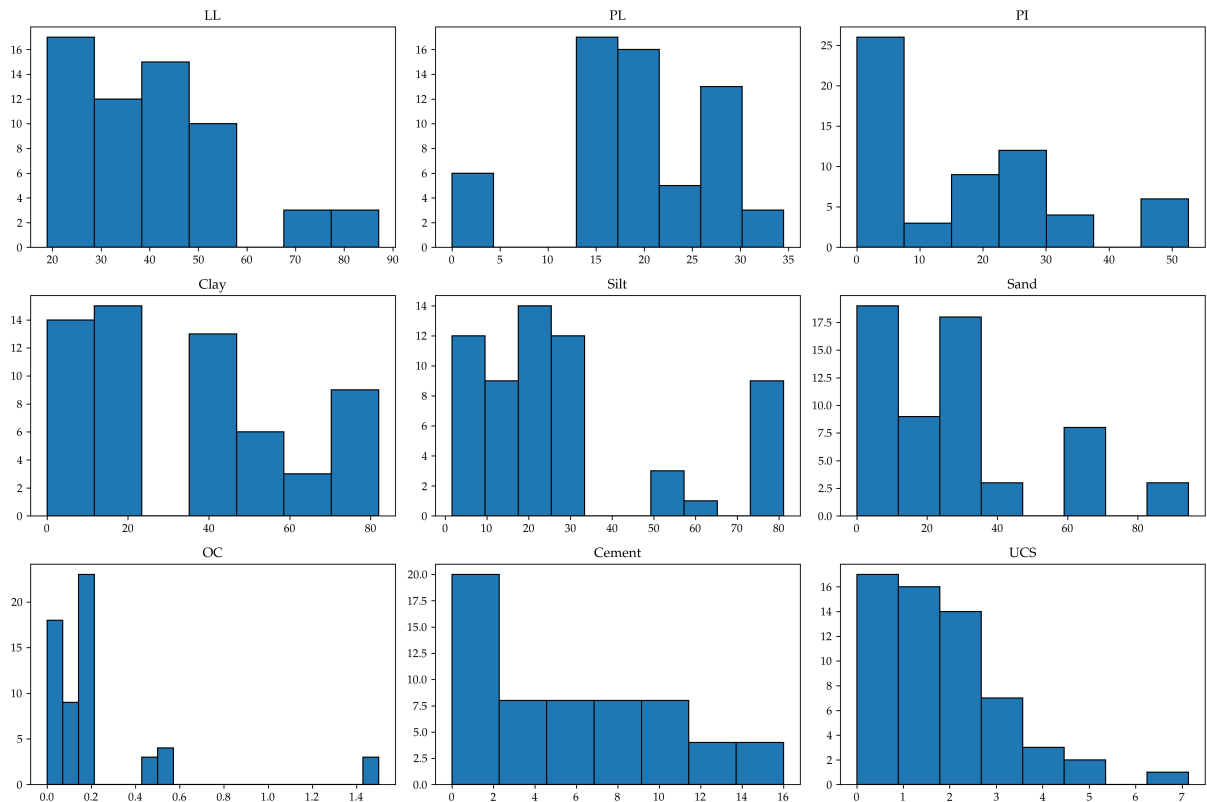
Fonte: Elaborado pelo autor (2026).

Tabela 4 – Estatísticas básicas do conjunto de dados D2.

Variável	Nome	Média	DP	Mín	25%	50%	75%	Máx
LL	Limite de liquidez (%)	40,1	18,2	18,9	24,6	42,4	51,0	87,1
PL	Limite de plasticidade (%)	19,6	8,4	0,0	16,4	20,0	26,0	34,5
PI	Índice de plasticidade (%)	18,3	15,9	0,0	5,0	21,9	30,0	52,6
Clay	Teor de argila (%)	33,7	24,9	0,0	15,0	38,5	47,8	82,0
Silt	Teor de silte (%)	29,4	25,0	1,6	13,8	22,3	30,0	81,1
Sand	Teor de areia (%)	27,3	26,3	0,0	5,5	23,9	31,5	94,4
OC	Teor de matéria orgânica (%)	0,2	0,3	0,0	0,0	0,2	0,2	1,5
Cement	Teor de cimento (%)	5,9	4,7	0,0	2,0	6,0	10,0	16,0
UCS	Resist. à compressão simples (MPa)	1,8	1,5	0,0	0,7	1,6	2,5	7,1

Fonte: Elaborado pelo autor (2025).

Figura 15 – Histogramas das variáveis do conjunto de dados D2.



Fonte: Elaborado pelo autor (2026).

Todas as amostras apresentam geometria cilíndrica, com altura e diâmetro de 50,8 mm, e foram coletadas em diferentes estados dos EUA, incluindo Iowa, Flórida, Illinois, Texas, Michigan, Virgínia, Carolina do Norte e Tennessee. Os valores de RCS foram obtidos a partir de corpos de prova curados por sete dias em condição saturada, com compactações realizadas com energia equivalente ao ensaio Proctor normal.

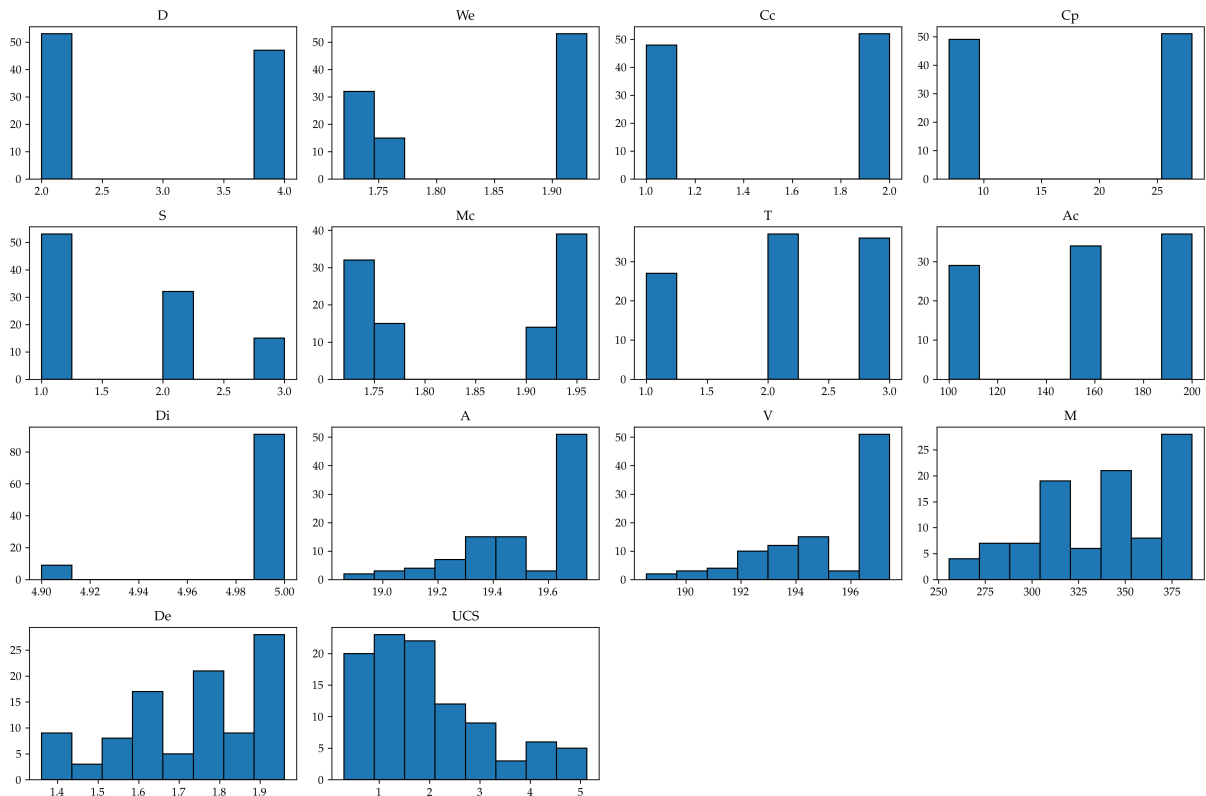
3.1.2 CONJUNTO DE DADOS D3

Tabela 5 – Estatísticas básicas do conjunto de dados D3.

Variável	Nome	Média	DP	Mín	25%	50%	75%	Máx
D	Profundidade de amostragem (m)	2,9	1,0	2,0	2,0	2,0	4,0	4,0
We	Densidade úmida (g/cm ³)	1,8	0,1	1,7	1,7	1,9	1,9	1,9
Cc	Condição de cura	1,5	0,5	1,0	1,0	2,0	2,0	2,0
Cp	Período de cura (dias)	17,7	10,6	7,0	7,0	28,0	28,0	28,0
S	Tipo de solo	1,6	0,7	1,0	1,0	1,0	2,0	3,0
Mc	Teor de umidade (%)	1,8	0,1	1,7	1,7	1,9	1,9	2,0
T	Tipo de cimento	2,1	0,8	1,0	1,0	2,0	3,0	3,0
Ac	Quantidade de cimento (kg/cm ³)	154,0	40,6	100,0	100,0	150,0	200,0	200,0
Di	Diâmetro do corpo de prova (cm)	5,0	0,0	4,9	5,0	5,0	5,0	5,0
A	Área do corpo de prova (cm ²)	19,5	0,2	18,9	19,3	19,6	19,6	19,7
V	Volume do corpo de prova (cm ³)	194,8	2,0	188,6	193,4	196,3	196,3	197,4
M	Massa do corpo de prova (g)	337,2	35,8	255,6	311,1	343,0	376,1	385,8
De	Densidade do corpo de prova (g/cm ³)	1,7	0,2	1,4	1,6	1,8	1,9	2,0
UCS	Resist. à compressão simples (MPa)	2,0	1,2	0,3	1,1	1,7	2,6	5,1

Cc: 1—ambiente interno, 2—ambiente externo; S: 1—argila amarela, 2—argila arenosa orgânica preta, 3—argila arenosa preta; T: 1—cimento Vissai, 2—cimento Nghi Son, 3—cimento Chinfon Hai Phong.
 Fonte: Elaborado pelo autor (2025).

Figura 16 – Histogramas das variáveis do conjunto de dados D3.



Fonte: Elaborado pelo autor (2026).

O conjunto de dados D3 utilizou informações experimentais apresentadas por Ngo et al. (49), contendo 14 variáveis de entrada, para treinar um modelo de aprendizado de máquina destinado à predição da RCS de misturas solo-cimento. As amostras foram coletadas na cidade de Hai Duong, no Vietnã. As variáveis podem ser agrupadas em três categorias. A primeira inclui características do solo e da mistura, como tipo de solo (S), teor de umidade (Mc), densidade úmida (We), profundidade de amostragem (D) e quantidade de cimento (Ac). A segunda categoria reúne propriedades geométricas e físicas dos corpos de prova após a mistura, como diâmetro (Di), área (A), volume (V), massa (M) e densidade (De). A terceira categoria contempla as condições de cura, incluindo condição de cura (Cc), período de cura (Cp) e tipo de cimento (T). Os tipos de cimento utilizados são encontrados no norte do Vietnã.

3.1.3 CONJUNTO DE DADOS D4

O conjunto D4 baseia-se nos resultados experimentais apresentados por Priyadarshee et al. (50) para uma argila caulínica. De acordo com o USCS, o solo foi classificado como CL (argila de baixa plasticidade). O material apresentou densidade relativa igual a 2,7; limite de liquidez de 43,3%; limite de plasticidade de 19,5% e índice de plasticidade de

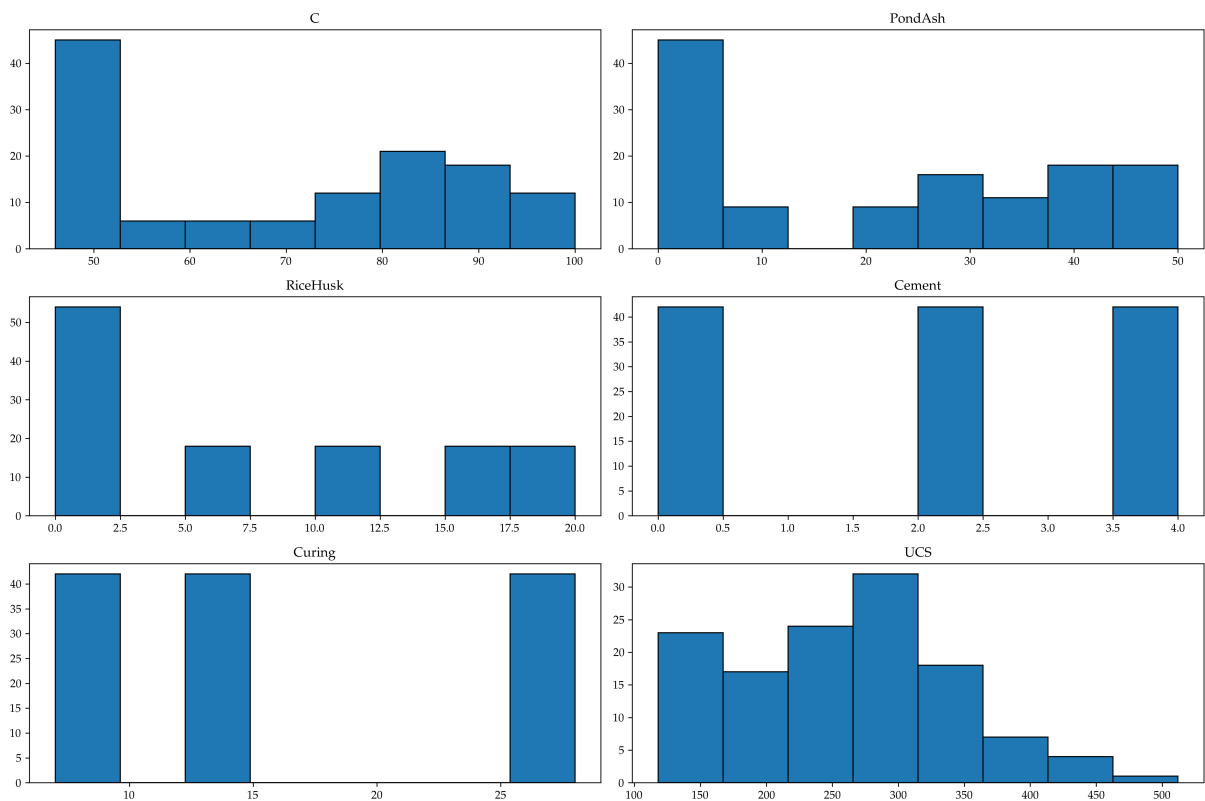
23,8. O teor ótimo de umidade obtido foi de 18,3%, e a massa específica seca máxima foi de 17,5 kN/m³.

Tabela 6 – Estatísticas básicas do conjunto de dados D4.

Variável	Nome	Média	DP	Mín	25%	50%	75%	Máx
C	Teor de argila (%)	69,4	18,7	46,0	50,0	73,0	86,0	100,0
PondAsh	Teor de cinza de carvão (%)	21,5	18,7	0,0	0,0	25,0	40,0	50,0
RiceHusk	Teor de cinza de casca de arroz (%)	7,1	7,5	0,0	0,0	5,0	15,0	20,0
Cement	Teor de cimento (%)	2,0	1,6	0,0	0,0	2,0	4,0	4,0
Curing	Período de cura (dias)	16,3	8,8	7,0	7,0	14,0	28,0	28,0
UCS	Resist. à compressão simples (kPa)	258,7	80,7	118,0	186,2	265,0	311,5	512,0

Fonte: Elaborado pelo autor (2025).

Figura 17 – Histogramas das variáveis do conjunto de dados D4.



Fonte: Elaborado pelo autor (2026).

Foram consideradas como variáveis de entrada o teor de argila (C), o teor de cinza de carvão (PondAsh), o teor de cinza de casca de arroz (RiceHusk), o teor de cimento

(Cement) e o período de cura (Curing). Essas variáveis foram utilizadas como preditores no desenvolvimento dos modelos, tendo a RCS como variável de saída. O estudo não informa o local exato de coleta das amostras, entretanto, considerando que a Índia possui numerosos depósitos de argila caulínica, um mineral amplamente extraído no país, e que os autores são afiliados a instituições indianas, é razoável inferir que os experimentos e a coleta de amostras foram provavelmente realizados na Índia.

A cinza de carvão utilizada como adição é um resíduo gerado pela queima de carvão em usinas termelétricas. As partículas apresentaram forma arredondada a subarredondada, coloração cinza-clara, densidade relativa de 2,10 e comportamento não plástico. A fração de silte e argila correspondeu a 41,6%, enquanto a fração de areia fina representou 44,6%.

A cinza de casca de arroz é um resíduo proveniente do beneficiamento do arroz. Aproximadamente 20–25% da massa do grão corresponde à casca, e cerca de 25% dessa casca é convertida em cinza após a queima. As propriedades do material dependem da temperatura de combustão. A cinza apresentou comportamento não plástico, sendo composta majoritariamente por SiO_2 , Fe_2O_3 e Al_2O_3 .

3.1.4 CONJUNTO DE DADOS D5

Tabela 7 – Estatísticas básicas do conjunto de dados D5.

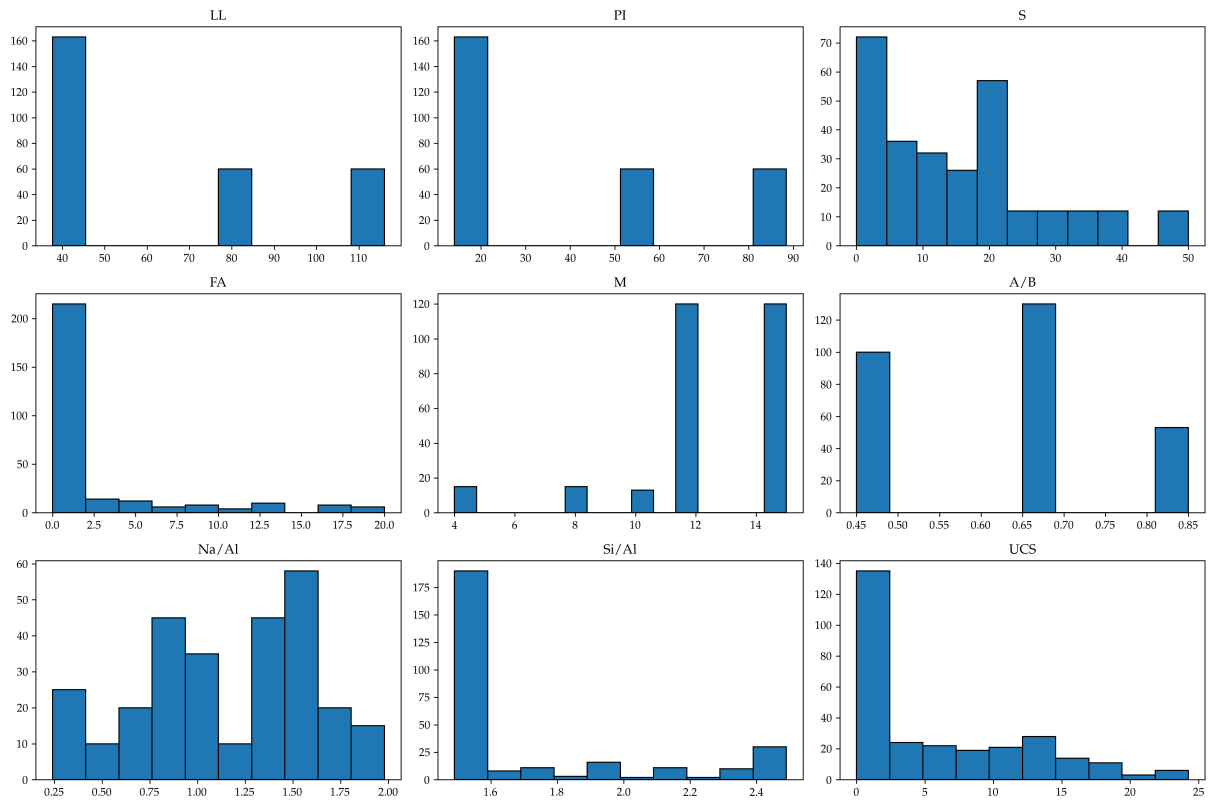
Variável	Nome	Média	DP	Mín	25%	50%	75%	Máx
LL	Limite de liquidez (%)	63,8	32,3	37,7	37,7	38,0	82,2	116,0
PI	Índice de plasticidade (%)	38,8	30,7	14,1	14,1	14,1	56,5	88,5
S	Teor de GGBS (%)	15,9	12,9	0,0	4,0	16,0	20,0	50,0
FA	Teor de cinza volante (%)	2,1	4,7	0,0	0,0	0,0	0,0	20,0
M	Concentração molar	12,4	2,7	4,0	12,0	12,0	14,5	15,0
A/B	Razão álcali/ligante	0,6	0,1	0,5	0,5	0,7	0,7	0,8
Na/Al	Razão Na/Al	1,2	0,4	0,2	0,9	1,2	1,5	2,0
Si/Al	Razão Si/Al	1,7	0,3	1,5	1,5	1,5	1,9	2,5
UCS	Resist. à compressão simples (MPa)	5,8	6,5	0,0	0,1	2,9	10,9	24,3

Na: sódio; Al: alumínio; Si: silício.

Fonte: Elaborado pelo autor (2025).

O conjunto de dados D5 foi extraído do estudo de Mozumder and Laskar (51), que investigou o uso de escória granulada de alto-forno moída (*Ground granulated blast furnace slag* - GGBS) e cinza volante. As variáveis de entrada incluem limites de Atterberg – limite de liquidez (LL) e índice de plasticidade (PI) –, porcentagem de GGBS (S), porcentagem de cinza volante (FA), concentração molar da solução alcalina (M), razão álcali-aglomerante (A/B), razão sódio-alumínio da solução alcalina (Na/Al) e razão silício-alumínio da solução alcalina (Si/Al). A variável de saída foi a RCS aos 28 dias, expressa em MPa.

Figura 18 – Histogramas das variáveis do conjunto de dados D5.



Fonte: Elaborado pelo autor (2026).

A solução alcalina utilizada neste estudo é formada por hidróxido de sódio (NaOH) com pureza de 98%, densidade específica de 2,13 e peso molecular de 40; e silicato de sódio (Na_2SiO_3) com densidade específica de 1,5, pureza de 97% e peso molecular igual a 212.

As amostras de solos foram coletadas com 2 metros de profundidade, na cidade de Silchar (Índia). Algumas variáveis não foram fornecidas na base de dados do estudo de Mozumder and Laskar (51) (descritas na Tabela 7), entretanto foram descritas no artigo e na Tabela 8.

Tabela 8 – Propriedades físicas e classificação dos solos do conjunto de dados D5.

Solo	LL	PL	PI	MDD (kN/m^3)	OMC (%)	d_{10} (μm)	d_{50} (μm)	d_{90} (μm)	USCS
S1	116,27	27,81	88,46	14,11	23,89	0,84	1,57	2,51	CH
S2	82,15	25,69	56,46	15,29	19,26	6,47	9,25	18,34	CH
S3	37,68	23,61	14,07	16,56	19,05	54,89	51,48	66,59	CL

LL: limite de liquidez; PL: limite de plasticidade; PI: índice de plasticidade; MDD: peso específico seco máximo; OMC: umidade ótima; d_{10} , d_{50} e d_{90} : diâmetros característicos da distribuição granulométrica. Fonte: Adaptado de Mozumder and Laskar (2015).

3.1.5 CONJUNTO DE DADOS D6

Tabela 9 – Estatísticas básicas do conjunto de dados D6.

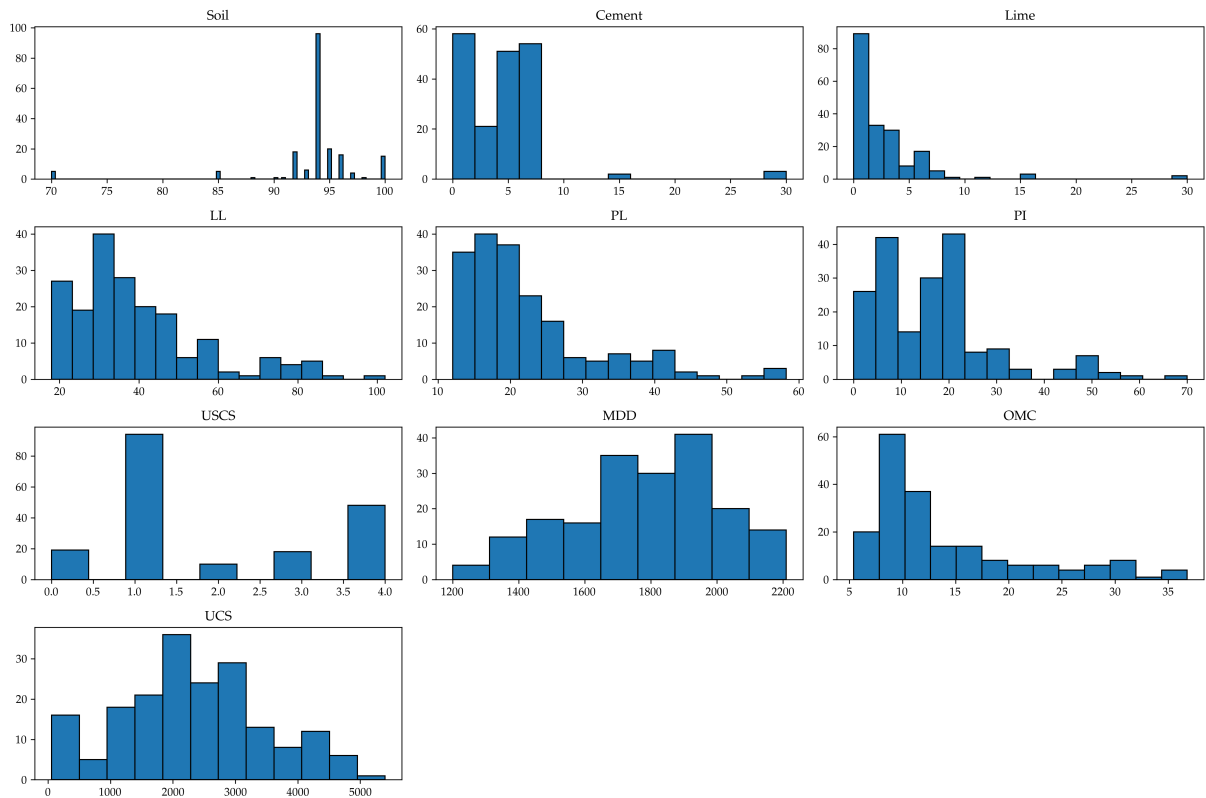
Variável	Nome	Média	DP	Mín	25%	50%	75%	Máx
Soil	Teor de solo (%)	93,7	4,7	70,0	94,0	94,0	95,0	100,0
Cement	Teor de cimento (%)	3,8	4,3	0,0	0,0	4,0	6,0	30,0
Lime	Teor de cal (%)	2,6	4,1	0,0	0,0	2,0	4,0	30,0
LL	Limite de liquidez (%)	39,5	16,7	18,0	29,0	35,0	45,0	102,0
PL	Limite de plasticidade (%)	22,7	9,4	12,0	17,0	20,0	26,0	58,2
PI	Índice de plasticidade (%)	16,8	12,7	0,0	7,0	15,5	21,0	70,0
USCS	Sistema Unificado de Classificação de Solos	1,9	1,4	0,0	1,0	1,0	4,0	4,0
MDD	Peso específico seco máximo (MN/m ³)*	1,8	0,2	1,2	1,6	1,8	2,0	2,2
OMC	Teor de umidade ótima (%)	13,9	7,3	5,4	8,8	11,2	16,3	36,8
UCS	Resist. à compressão simples (MPa)*	2,3	1,2	0,1	1,7	2,3	3,0	5,4

O teor de solo, cimento e cal é expresso como porcentagem em massa da mistura total; Tipos de solo empregados (USCS): argila de baixa plasticidade (CL), argila de alta plasticidade (CH), silte de baixa plasticidade (ML), silte de alta plasticidade (MH) e silte argiloso de baixa plasticidade (CL-ML); (*) Os valores de MDD e UCS utilizados para o treinamento do modelo foram originalmente expressos em kN/m³ e kPa, respectivamente. A conversão de unidades nesta tabela foi adotada apenas para fins de apresentação. Fonte: Elaborado pelo autor (2025).

O conjunto D6 apresentado por Taffese and Abegaz (3) reúne dados experimentais de dez bases de dados pública, e teve como objetivo a predição da RCS a partir de solos estabilizados com cimento e cal. As variáveis de entrada descrevem proporções das misturas e tipos de solos estabilizados – teor de solo (Soil), teor de cimento (Cement) e teor de cal (Lime) –, limites de Atterberg – limite de liquidez (LL), limite de plasticidade (PL) e índice de plasticidade (PI) –, índices de classificação dos solos (USCS), e parâmetros de compactação, como umidade ótima (OMC) e peso específico seco máximo (MDD). A variável de saída foi a RCS aos 28 dias, expressa em MPa.

O conjunto contempla solos de doze países, incluindo regiões da África (Chlef, Argélia), Américas (Oklahoma, EUA), Oriente Médio (Trabzon, Turquia), Sul da Ásia (Índia e Bangladesh) e Oceania (Queensland e Nova Gales do Sul, Austrália). (104, 105, 106, 107, 108, 109, 110, 111, 112, 113).

Figura 19 – Histogramas das variáveis do conjunto de dados D6.



Fonte: Elaborado pelo autor (2026).

3.1.6 CONJUNTO DE DADOS D7

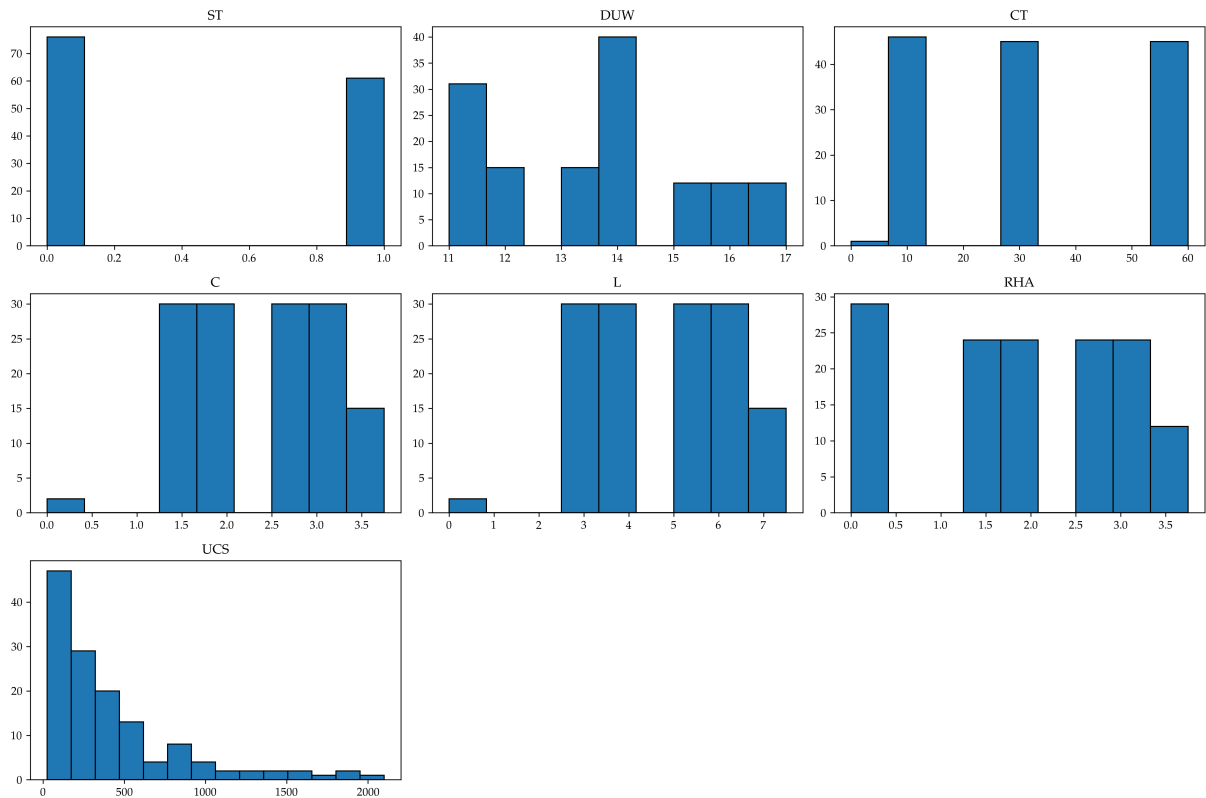
Tabela 10 – Estatísticas básicas do conjunto de dados D7.

Variável	Nome	Média	DP	Mín	25%	50%	75%	Máx
ST	Tipo de solo	0,4	0,5	0,0	0,0	0,0	1,0	1,0
DUW	Peso específico seco (kN/m ³)	13,5	1,9	11,0	12,0	14,0	15,0	17,0
CT	Tempo de cura (dias)	31,3	22,0	0,0	7,0	28,0	60,0	60,0
C	Teor de cimento (%)	2,3	0,9	0,0	1,9	2,5	3,1	3,8
L	Teor de cal (%)	4,7	1,7	0,0	3,8	5,0	6,2	7,5
RHA	Teor de cinza de casca de arroz (%)	1,9	1,2	0,0	1,2	1,9	3,1	3,8
UCS	Resist. à compressão simples (kPa)	432,9	430,9	25,0	138,0	285,0	539,0	2099,0

ST: SM—areia siltosa; MH—silte de alta plasticidade.

Fonte: Elaborado pelo autor (2025).

Figura 20 – Histogramas das variáveis do conjunto de dados D7.



Fonte: Elaborado pelo autor (2026).

O conjunto D7 foi elaborado com base no estudo de Tabarsa et al. (52), que avaliou a RCS de areia siltosa e silte de alta plasticidade, estabilizados com diferentes combinações de cimento, cal e cinza de casca de arroz, na Malásia. As variáveis de entrada utilizadas no treinamento dos modelos foram tipo de solo (ST), peso específico seco (DUW), tempo de cura (CT), teor de cimento (C), teor de cal (L) e teor de cinza de casca de arroz (RHA).

Algumas variáveis não foram fornecidas na base de dados do estudo de Tabarsa et al. (52) (descritas na Tabela 10), entretanto foram descritas no artigo e na Tabela 11.

Tabela 11 – Propriedades físicas e classificação dos solos do conjunto de dados D7.

Propriedade	Solo SM	Solo MH	Método de ensaio
Teor de areia (%)	51,70	43,52	–
Teor de silte (%)	28,60	33,87	–
Teor de argila (%)	7,80	22,61	–
Teor de matéria orgânica (%)	1,20	12,80	–
Densidade relativa	2,55	2,38	ASTM D854
Classificação pelo USCS	Areia siltosa	Silte de alta plasticidade	ASTM D2487
Limite de liquidez (%)	48,10	55,40	ASTM D4318
Limite de plasticidade (%)	31,50	36,40	ASTM D4318
Índice de plasticidade (%)	16,60	19,00	ASTM D4318
Umidade ótima (%)	16,30	24,65	ASTM D698
Peso específico seco máximo (kN/m ³)	17,52	14,67	ASTM D698
Potencial hidrogeniônico (pH)	4,78	4,21	ASTM D4972

Fonte: Adaptado de Tabarsa et al. (2021).

3.1.7 CONJUNTO DE DADOS D8

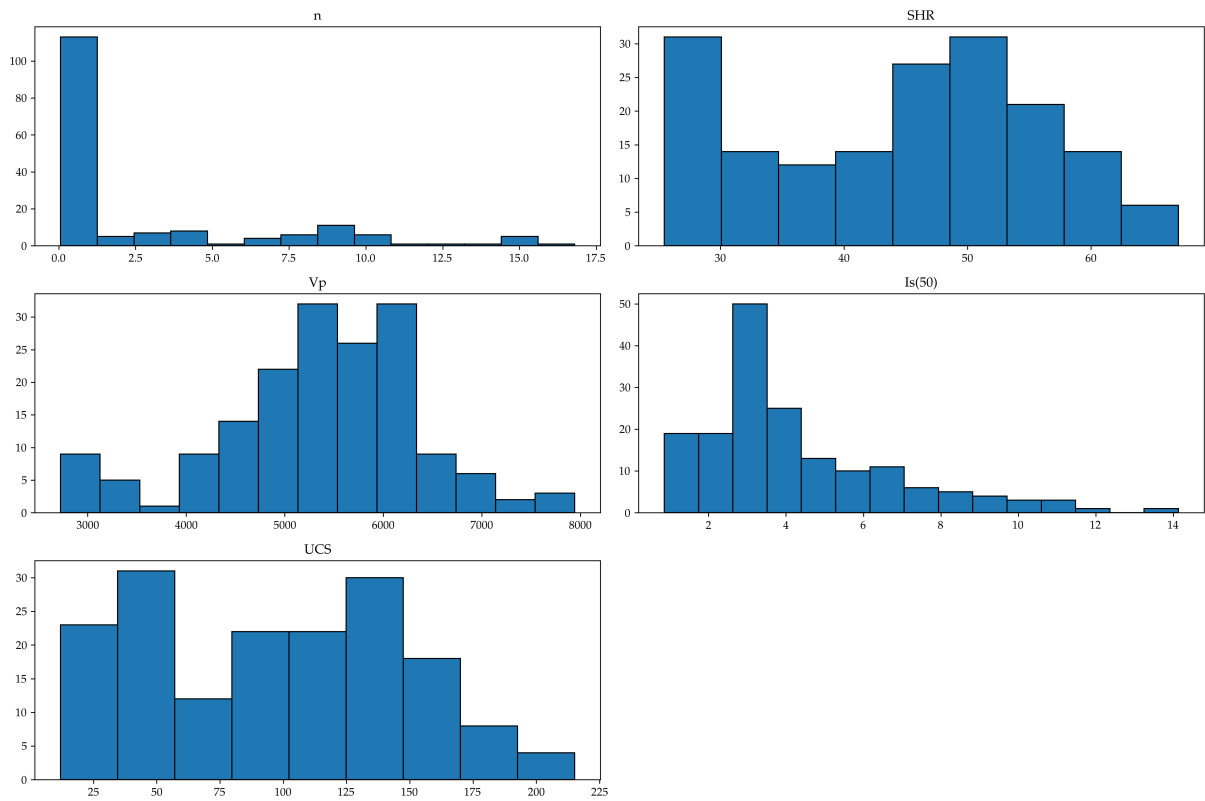
O estudo de Mahmoodzadeh et al. (53) utilizou o conjunto D8 para a predição da RCS de diferentes tipos de rochas, como argilito, granito, xisto, arenito, travertino, calcário, ardósia, dolomito e marga, localizados no Irã. Os modelos de predição empregaram como variáveis de entrada a porosidade (n), o índice de rebote do martelo de Schmidt (SH), a velocidade da onda P (V_p) e o índice de carga pontual ($I_{s(50)}$). Este conjunto de rochas foi apresentado de forma isolada em relação aos demais compostos por solos, com o objetivo de verificar o comportamento dos sistemas, ou seja, se estes conseguem manter o mesmo nível de desempenho.

Tabela 12 – Estatísticas básicas do conjunto de dados D8.

Variável	Nome	Média	DP	Mín	25%	50%	75%	Máx
n	Porosidade (%)	2,8	4,1	0,1	0,3	0,5	3,8	16,8
SHR	Número do martelo de Schmidt	44,4	11,5	25,5	33,5	46,0	53,0	67,1
Vp	Velocidade da onda P (km/s) ^(*)	5,4	1,0	2,7	4,9	5,5	6,0	7,9
Is ₍₅₀₎	Índice de carga pontual (MPa)	4,3	2,5	0,9	2,8	3,4	5,4	14,1
UCS	Resist. à compressão simples (MPa)	96,0	51,1	12,0	41,6	99,5	136,7	215,2

(*) Os valores de Vp utilizados para o treinamento do modelo foram originalmente expressos em m/s. A conversão de unidades nesta tabela foi adotada apenas para fins de apresentação. Fonte: Elaborado pelo autor (2025).

Figura 21 – Histogramas das variáveis do conjunto de dados D8.



Fonte: Elaborado pelo autor (2026).

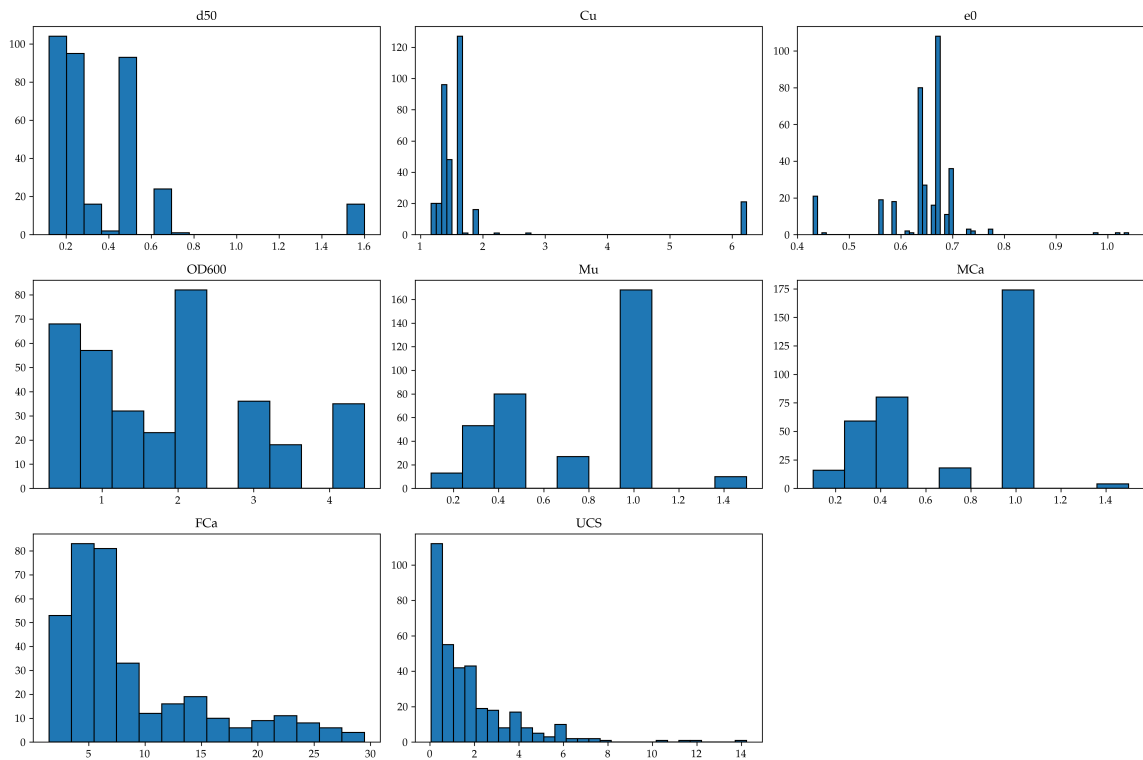
3.1.8 CONJUNTO DE DADOS D9

Tabela 13 – Estatísticas básicas do conjunto de dados D9.

Variável	Nome	Média	DP	Mín	25%	50%	75%	Máx
d50	Diâmetro médio dos grãos (mm)	0,4	0,3	0,1	0,2	0,3	0,5	1,6
Cu	Coefficiente de uniformidade	1,8	1,1	1,2	1,4	1,4	1,6	6,2
e0	Índice de vazios inicial	0,6	0,1	0,4	0,6	0,7	0,7	1,0
OD600	Turbidez (OD ₆₀₀)	2,0	1,2	0,3	1,0	1,8	3,0	4,5
Mu	Concentração de ureia (mol/L)	0,7	0,3	0,1	0,5	1,0	1,0	1,5
MCa	Concentração de cálcio (mol/L)	0,7	0,3	0,1	0,5	1,0	1,0	1,5
FCa	Teor de carbonato de cálcio (%)	8,8	6,5	1,5	4,2	6,5	11,7	29,5
UCS	Resist. à compressão simples (MPa)	1,8	2,0	0,1	0,5	1,2	2,3	14,2

Fonte: Elaborado pelo autor (2025).

Figura 22 – Histogramas das variáveis do conjunto de dados D9.



Fonte: Elaborado pelo autor (2026).

O conjunto de dados D9 foi construído a partir da compilação de treze estudos disponíveis na literatura, conforme apresentado por Wang and Yin (20), com o objetivo de prever a RCS de solos estabilizados por meio de precipitação de calcita induzida por microrganismos (MICP). Essa técnica utiliza bactérias produtoras de urease que exercem a função de catalisadoras para obtenção de solução ligante. Todas as amostras de solos foram estabilizadas com 100% de saturação. Sete variáveis foram consideradas no treinamento dos modelos: diâmetro médio das partículas (d_{50}), coeficiente de uniformidade (C_u), índice de vazios inicial (e_0), turbidez da solução bacteriana (OD600), concentração de ureia (M_u), concentração de cálcio (M_{Ca}) e quantidade de carbonato de cálcio (FCa).

As amostras foram obtidas em diferentes regiões, incluindo Alemanha, Austrália, EUA (Ottawa, Mississippi e Nevada) e China (114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125).

3.1.9 CONJUNTO DE DADOS D10

Por fim, o conjunto D10 foi utilizado no estudo de Zhang et al. (9), cujo objetivo foi a predição da RCS de solos estabilizados com cimento. O modelo foi treinado a partir das variáveis proporção de cimento (C_c), teor de água (C_w) e período de cura (T_c). Como o estudo visa a utilização de solo-cimento com barras de polímero reforçadas com fibra de vidro (GFRP) para aplicações em obras de engenharia, também foram avaliadas as variáveis relacionadas ao ensaio de aderência à tração entre o solo-cimento e as barras GFRP: resistência de aderência à tração máxima (T_p) e resistência de aderência à tração residual (T_t).

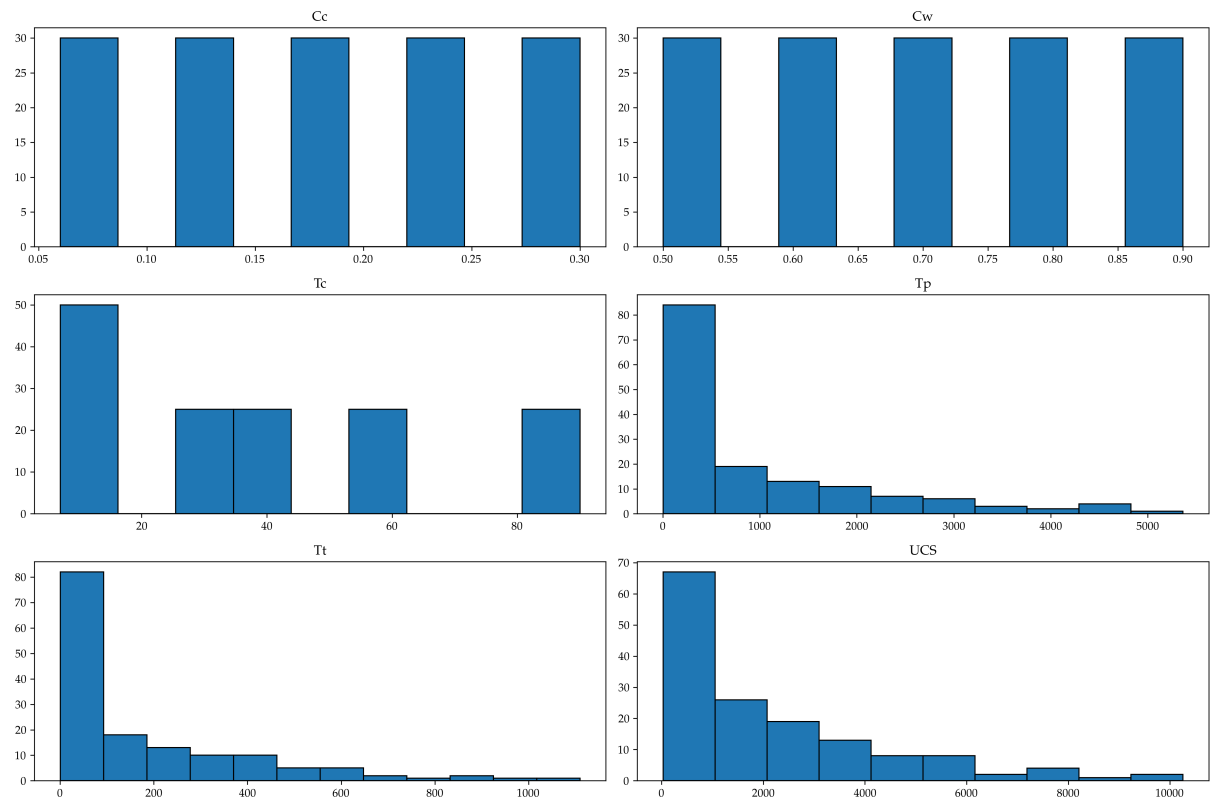
Tabela 14 – Estatísticas básicas do conjunto de dados D10.

Variável	Nome	Média	DP	Mín	25%	50%	75%	Máx
C_c	Proporção de cimento (kg/m^3)	0,2	0,1	0,09	0,1	0,2	0,2	0,3
C_w	Teor de água (kg/m^3)	0,7	0,1	0,49	0,6	0,7	0,8	0,9
T_c	Período de cura (dias)	40,2	28,4	7,00	14,0	35,0	60,0	90,0
T_p	Resist. de aderência máxima (kPa)	954,1	1173,6	3,00	95,5	419,2	1524,0	5363,1
T_t	Resist. de aderência residual (kPa)	179,5	225,3	0,29	18,7	74,6	276,3	1109,6
UCS	Resist. à compressão simples (MPa) ^(*)	2,1	2,2	0,03	0,4	1,3	3,2	10,3

(*) Os valores de UCS utilizados para o treinamento do modelo foram originalmente expressos em kPa. As conversões de unidades nesta tabela foram adotadas exclusivamente para fins de apresentação. Fonte: Elaborado pelo autor (2025).

As amostras de solo foram coletadas na China, da planície de inundação na confluência dos rios Xiangjiang e Jinjiang em Changsha. Foi seco ao ar, moído em uma máquina e, em seguida, peneirado. Apenas partículas com diâmetro inferior a 5 mm foram utilizadas neste experimento. Possui densidade relativa de 2,705; teor de umidade natural de 30 a 90%; Limite de liquidez igual a 58,1% e Limite de plasticidade igual a 28,6%.

Figura 23 – Histogramas das variáveis do conjunto de dados D10.



Fonte: Elaborado pelo autor (2026).

3.1.10 COMPARAÇÃO DOS CONJUNTOS DE DADOS

A Figura 24 resume, dos principais parâmetros que caracterizam solos, quais tipos de informações cada conjunto de dados possui. Como o conjunto D8 trata de amostras de rochas, o mesmo não foi considerado.

Note que informações sobre o estabilizante aparecem em todos os conjuntos de dados afinal, é escopo do presente estudo a análise de solos estabilizados. Informações a respeito do tipo de solo ou da composição do solo utilizado aparecem com alta frequência, sendo exceção apenas os conjuntos D5 e D10; entretanto, observe que o conjunto D5 apresenta informações a respeito dos limites de Atterberg, que podem ser utilizados como forma de classificar os solos.

Os limites de Atterberg são informados nos conjuntos D1, D2, D5 e D6, enquanto informações sobre a cura das amostras aparecem nos conjuntos D3, D4, D7 e D10.

Informações sobre a compactação das amostras aparecem nos conjuntos D6 e D7. O conjunto D3 foi o único que trouxe informações sobre a densidade das amostras. Os conjuntos D3, D6 e D7 se destacam por apresentarem muitas informações relacionadas com parâmetros de caracterização dos solos.

Figura 24 – Informações disponíveis nos conjuntos de dados analisados.

Conjunto	Estabilizante	Tipo de solo	Limites de Atterberg	Densidade	Cura	Compactação
D1	x	x	x			
D2	x	x	x			
D3	x	x		x	x	
D4	x	x			x	
D5	x		x			
D6	x	x	x			x
D7	x	x			x	x
D8	-	-	-	-	-	-
D9	x	x				
D10	x				x	

Fonte: Elaborado pelo autor (2026).

3.2 DESENVOLVIMENTO E VALIDAÇÃO DOS SISTEMAS

A escolha dos sistemas de aprendizado de máquina automatizado considerados neste estudo seguem alguns critérios práticos. Foram selecionadas apenas ferramentas abertas e gratuitas, o que permite total reprodutibilidade dos experimentos. Além disso, optou-se por sistemas que adotam estratégias distintas de busca e ajuste, pois isso amplia a diversidade dos modelos avaliados. Outro ponto é que todos os sistemas escolhidos já aparecem de forma consolidada na literatura de AutoML e contam com comunidades ativas de desenvolvimento (82). Por fim, todos oferecem suporte direto para tarefas de regressão em dados tabulares, o que atende ao problema tratado aqui, relacionado à predição da RCS. Soluções proprietárias ou dependentes de serviços em nuvem não foram consideradas para evitar limitações externas e garantir condições de avaliação totalmente reproduzíveis.

Os sistemas de aprendizado de máquina automatizado avaliados neste trabalho adotam estratégias distintas para o tratamento de valores ausentes e imputação de dados. Essas diferenças podem influenciar o comportamento dos modelos, sobretudo em conjuntos de dados pequenos. Neste estudo nenhuma imputação foi aplicada, pois não haviam valores ausentes após a etapa de integração dos dados.

As tarefas de pré-processamento dos dados, como normalização, tratamento de variáveis categóricas e identificação de valores atípicos, foram executadas de forma nativa

por cada sistema. Para representar cenários de aplicação mais próximos da prática, utilizaram-se as configurações padrão recomendadas pelas próprias ferramentas.

A definição dos hiperparâmetros é uma etapa fundamental para permitir a reprodutibilidade, a comparabilidade e a imparcialidade de experimentos que envolvem diferentes sistemas de AutoML. Neste trabalho, a configuração de cada sistema foi definida com base em sua documentação oficial e ajustada de acordo com o recurso computacional disponível (tempo de execução) e as características dos conjuntos de dados utilizados.

Os parâmetros listados na Tabela 15 permaneceram fixos ao longo de todos os experimentos, de modo que as comparações entre os sistemas foram realizadas sob as mesmas restrições computacionais. O limite de tempo foi escolhido com base em testes preliminares, que indicaram que 120 segundos (o limite global para cada execução) foram suficientes para encontrar resultados ótimos ou próximos do ótimo para os modelos, equilibrando o objetivo de obter o melhor desempenho preditivo com a necessidade de eficiência computacional e tempo de desenvolvimento. Vale destacar que o sistema AutoKeras é o único que não possui parâmetro para limitar o tempo de execução.

Tabela 15 – Hiperparâmetros definidos para cada sistema de AutoML.

Sistema	Parâmetro	Valor
AutoGluon	label	target
	verbosity	False
	time_limit	120 (segundos)
AutoKeras	max_trials	50
	epochs	100
	column_names	feature_names
	loss	mean_absolute_error
	time_budget	120 (segundos)
FLAML	metric	MAE
	task	regression
	estimator_list	[lgbm, rf, xgboost, extra_tree, xgb_limitdepth, catboost, kneighbor]
	verbose	False
	n_jobs	2
H2O	nthreads	2
	max_runtime_secs	120 (segundos)
	sort_metric	RMSE
TPOT	max_time_mins	2 (minutos)
	generations	20
	population_size	20
	cv	5
	verbosity	False
	n_jobs	2

Fonte: Elaborado pelo autor (2025).

Abaixo segue um resumo com as descrições dos parâmetros informados na Tabela 15 para cada sistema:

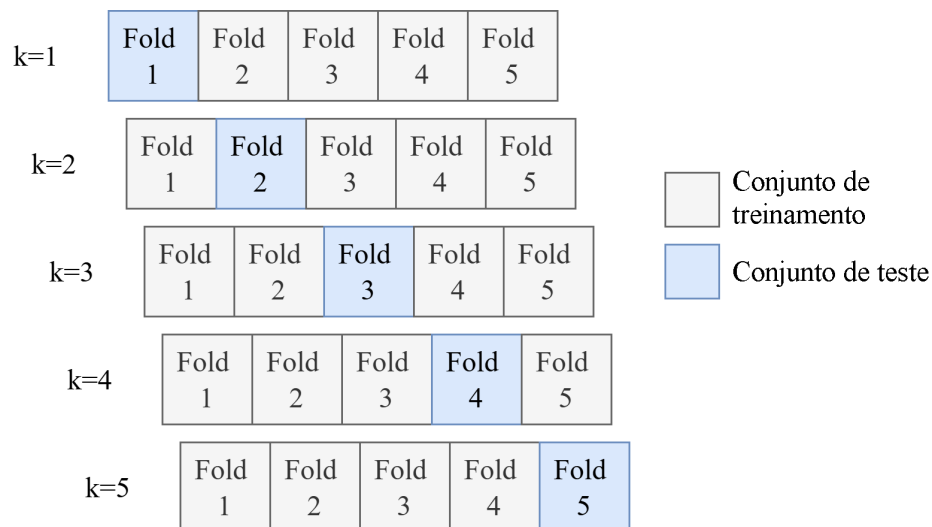
- AutoGluon - label: define a variável alvo do modelo;
- AutoGluon - verbosity: controla o nível de detalhamento de exibição de mensagens durante a execução;
- AutoGluon - time_limit: tempo máximo permitido para o treinamento e ajuste automático dos modelos;
- AutoKeras - max_trials: número máximo de arquiteturas/modelos avaliados durante a busca automática;
- AutoKeras - epochs: número máximo de épocas de treinamento para cada modelo;
- AutoKeras - column_names: define os nomes das variáveis de entrada;
- AutoKeras - loss: função de perda utilizada para otimização do modelo;
- FLAML - time_budget: tempo máximo permitido para o treinamento e ajuste automático dos modelos;
- FLAML - metric: métrica de desempenho utilizada para avaliação dos modelos;
- FLAML - task: tipo de tarefa de aprendizado de máquina;
- FLAML - estimator_list: lista de algoritmos candidatos considerados no processo de AutoML;
- FLAML - verbose: controla o nível de detalhamento de exibição de mensagens durante a execução;
- FLAML - n_jobs: número de núcleos utilizados para processamento paralelo;
- H2O - nthreads: número de núcleos utilizados para processamento paralelo;
- H2O - max_runtime_secs: tempo máximo permitido para o treinamento e ajuste automático dos modelos;
- H2O - sort_metric: métrica de desempenho utilizada para avaliação dos modelos;
- TPOT - max_time_mins: tempo máximo permitido para o treinamento e ajuste automático dos modelos;
- TPOT - generations: número de gerações do algoritmo genético;
- TPOT - population_size: número de candidatos em cada geração;

- TPOT - cv: número de subconjuntos utilizados na validação cruzada;
- TPOT - verbosity: controla o nível de detalhamento de exibição de mensagens durante a execução;
- TPOT - n_jobs: número de núcleos utilizados para processamento paralelo.

Devido às diferenças significativas entre os conjuntos de dados D1 a D10, incluindo variações nos materiais, nos procedimentos laboratoriais, nas variáveis preditoras disponíveis e no número de amostras (ver Seção 3.1), os modelos de AutoML foram treinados e avaliados de forma independente para cada conjunto. Não foi ajustado um único modelo com os dados combinados, em vez disso, o desempenho foi analisado separadamente entre os diferentes conjuntos de dados.

Cada sistema foi treinado e validado a partir de uma divisão estratificada dos conjuntos de dados, com 70% das amostras destinadas ao treinamento e à validação interna, e 30% reservadas para a etapa de teste. Todos os experimentos foram executados no mesmo ambiente computacional, considerando *hardware* e *software*, de forma a permitir uma comparação consistente entre os sistemas avaliados. Os resultados apresentados correspondem à média de 30 execuções independentes, o que aumenta a confiabilidade estatística das análises. Para a validação cruzada no conjunto de treinamento, adotou-se uma semente aleatória distinta em cada execução (*run*), definida por $seed = run \times 37 + 1001$. Esse procedimento assegura reprodutibilidade e variação controlada nas divisões dos dados.

Figura 25 – Exemplo de validação cruzada *k-fold* com cinco subconjuntos (*folds*).



Fonte: Elaborado pelo autor (2025).

A validação cruzada é utilizada para avaliar a capacidade preditiva dos modelos e reduzir erros associados ao *overfitting*. Neste trabalho, adotou-se a técnica de validação

cruzada k -fold, na qual o conjunto de dados é particionado em k subconjuntos de tamanhos semelhantes. Em cada iteração, o modelo é ajustado utilizando $k - 1$ subconjuntos e avaliado no subconjunto restante. Esse procedimento é repetido k vezes, de modo que cada subconjunto seja utilizado uma única vez como conjunto de validação (126). A Figura 25 ilustra o funcionamento da validação cruzada k -fold com cinco partições ($k = 5$): em cada rodada um dos subconjuntos é definido como conjunto de teste (em azul), enquanto os demais compõem o conjunto de treinamento (em cinza).

3.3 MÉTRICAS DE DESEMPENHO

A avaliação dos modelos foi baseada em métricas padrão de regressão, incluindo o Erro Absoluto Médio (*Mean Absolute Error* - MAE), o Erro Percentual Absoluto Médio (*Mean Absolute Percentage Error* - MAPE), a Raiz do Erro Quadrático Médio (*Root Mean Squared Error* - RMSE), o coeficiente de correlação de Pearson (R), e o coeficiente de determinação (R^2). Essas métricas foram selecionadas com o objetivo de analisar o desempenho dos sistemas sob diferentes perspectivas. Também foi utilizado o desvio padrão (DP) para avaliar a variância dos resultados. (Figura 26).

Figura 26 – Siglas e expressões das métricas de desempenho e do desvio padrão.

Sigla	Expressão
DP	$\sqrt{\frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2}$
MAE	$\frac{1}{N} \sum_{t=1}^N y_i - \hat{y}_i $
MAPE	$\frac{1}{N} \sum_{i=1}^N \left \frac{y_i - \hat{y}_i}{y_i} \right \times 100\%$
RMSE	$\frac{1}{N} \sqrt{\sum_{i=1}^N (y_i - \hat{y}_i)^2}$
R	$\frac{cov(y, \hat{y})}{\sqrt{var(y)} \sqrt{var(\hat{y})}}$
R^2	$\sum_{i=1}^N \frac{(y_i - \hat{y}_i)^2}{(y_i - \bar{y})^2}$

Fonte: Elaborado pelo autor (2025).

As métricas MAE e RMSE quantificam a magnitude dos erros de predição. O MAE fornece uma medida linear do erro médio, enquanto o RMSE é mais sensível a erros

elevados devido à penalização quadrática. O MAPE expressa o erro de forma percentual, facilitando a interpretação do desvio relativo no entanto, essa métrica pode apresentar instabilidade quando os valores reais se aproximam de zero. O coeficiente R indica a intensidade da relação linear entre os valores preditos e observados, enquanto o coeficiente de determinação R^2 representa a fração da variância explicada pelo modelo. Em conjunto, essas métricas possibilitam uma avaliação equilibrada da acurácia absoluta e relativa das predições, além do ajuste global do modelo.

Para permitir a comparação global do desempenho dos modelos a partir de diferentes métricas de regressão, foi definido um Índice de Desempenho composto (ID). Esse índice reúne, em um único valor, as informações fornecidas pelos indicadores apresentados na Tabela 26. A formulação do ID (Equação 3.1) considera simultaneamente medidas associadas à capacidade de correlação e à magnitude dos erros, produzindo uma métrica sintética em que valores mais altos indicam melhor qualidade preditiva.

$$\begin{aligned} \text{ID} = & w_1 \cdot R + w_2 \cdot R^2 + w_3 \cdot \left(1 - \frac{\text{RMSE}}{\text{RMSE}_{\max}}\right) \\ & + w_4 \cdot \left(1 - \frac{\text{MAE}}{\text{MAE}_{\max}}\right) + w_5 \cdot \left(1 - \frac{\text{MAPE}}{\text{MAPE}_{\max}}\right) \end{aligned} \quad (3.1)$$

Nessa formulação, w_i corresponde ao peso atribuído a cada métrica, enquanto "métrica_{max}" representa o maior valor observado dessa métrica entre todos os modelos avaliados em um determinado conjunto de dados, sendo interpretado como a pior condição relativa. Para assegurar uma contribuição equivalente de cada indicador e evitar viés na agregação, foi utilizada média aritmética simples, ou seja, todos os pesos foram definidos como 0,20.

3.4 IMPORTÂNCIA DOS ATRIBUTOS

A identificação dos atributos que mais influenciam as predições dos modelos é fundamental para avaliar a confiabilidade e a interpretabilidade dos sistemas de aprendizado de máquina. A análise de importância das variáveis permite verificar se os algoritmos priorizam entradas com significado físico ou geotécnico, o que contribui para a compreensão dos resultados obtidos.

Neste trabalho, a importância dos atributos foi avaliada a partir da combinação entre a correlação das variáveis com a variável alvo e as pontuações de importância por permutação. Essa abordagem fornece uma visão mais completa sobre como cada sistema de AutoML hierarquiza as variáveis de entrada.

Para avaliar a contribuição individual de cada variável nas predições, foi adotado o método de importância por permutação (127). Nesse procedimento, a relevância de

uma variável é estimada a partir do aumento do erro de predição quando seus valores são misturados aleatoriamente, quebrando sua relação com a variável resposta. Inicialmente, calcula-se um valor de referência de desempenho no conjunto de teste original, utilizando o erro quadrático médio como métrica. Em seguida, os valores de cada variável são permutados cem vezes, e a variação média no desempenho é utilizada como medida de importância. Diferenças médias maiores indicam maior influência da variável no modelo (127, 128).

Esse procedimento foi aplicado a todos os sistemas de AutoML avaliados, de forma a atender às exigências específicas de entrada de dados dessas ferramentas. Os resultados obtidos permitiram identificar as variáveis mais relevantes para o desempenho dos modelos e auxiliaram na interpretação da acurácia preditiva observada nos diferentes conjuntos de dados.

3.5 CONFIGURAÇÕES COMPUTACIONAIS

Os experimentos computacionais foram realizados em linguagem Python (versão 3.9.23) e ambiente virtualizado com Poetry (versão 1.8.2). O equipamento utilizado possui processador Intel(R) Core(TM) i5-4210U, com quatro núcleos a 1,70 GHz e 8 GB de memória RAM, operando sob o sistema Ubuntu Linux 24.04.3 LTS. A Tabela 16 apresenta as bibliotecas empregadas nos experimentos, bem como as respectivas versões utilizadas.

Tabela 16 – Bibliotecas e versões utilizadas.

Biblioteca	Versão
AutoGluon	0.7.0
AutoKeras	1.0.20
CMake	3.28.3
Cython	3.1.3
FLAML	1.2.4
g++	13.3.0
H2O	3.44.0.3
Jinja2	3.1.6
Joblib	1.5.1
Keras-Tuner	1.4.7
Matplotlib	3.9.4
NumPy	1.23.5
Openpyxl	3.1.2
Pandas	1.5.3
Permetrics	2.0.0
Scikit-Learn	1.2.2
Scipy	1.11.4
Seaborn	0.12.2
Skillmetrics	1.1.8
TPOT	0.12.1
Tensorflow	2.9.3

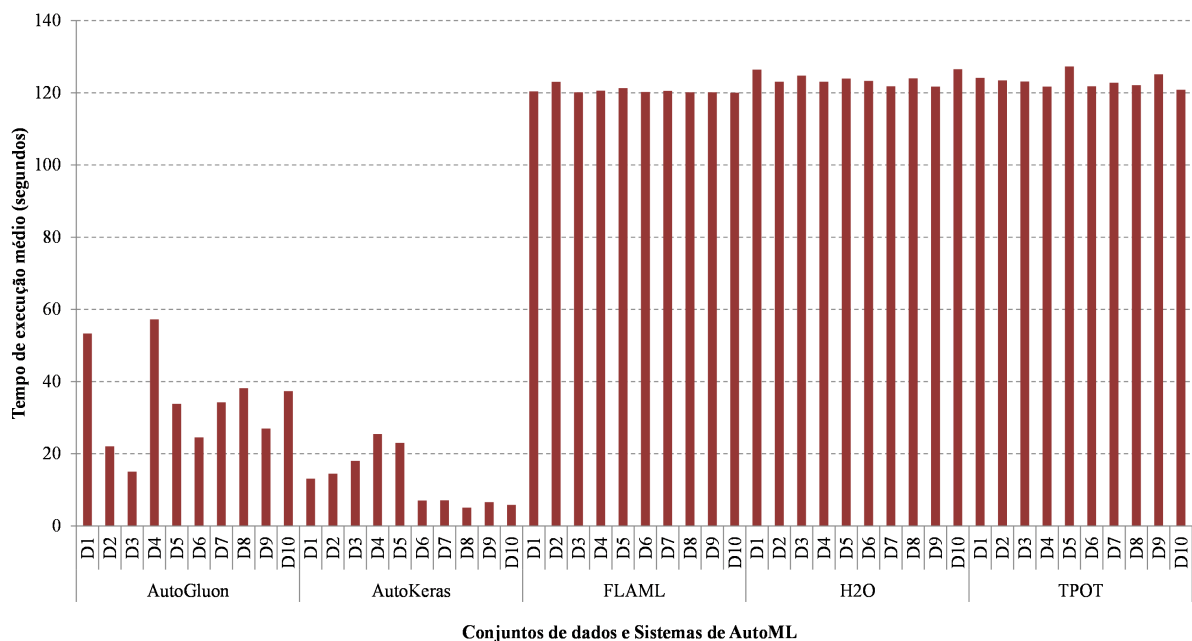
Fonte: Elaborado pelo autor (2025).

4 RESULTADOS

Os experimentos foram obtidos para cinco sistemas AutoML avaliados em dez conjuntos de dados (ambos detalhados na seção 3). Os resultados decorrem de trinta execuções repetidas e independentes, o que permite considerar a variabilidade inerente ao processo de ajuste dos sistemas.

A Figura 27 sintetiza o custo computacional médio dos experimentos, expresso em segundos, para cada combinação de conjunto de dados e sistema AutoML. Observa-se que AutoGluon e AutoKeras demandaram menor tempo de execução na maioria dos casos. Em seguida, situam-se os demais sistemas. Esse comportamento está diretamente associado às restrições impostas durante os experimentos. Não foram definidos limites explícitos de tempo para o AutoKeras, pois o mesmo não disponibiliza um mecanismo nativo para controle da duração da execução. Os demais sistemas foram configurados com um tempo máximo de execução de 120 segundos. Note que, mesmo com a duração máxima de execução definida, o sistema AutoGluon obteve tempos de convergência inferiores, conforme ilustrado na Figura 27.

Figura 27 – Tempo médio de execução para cada conjunto de dados e sistema AutoML.

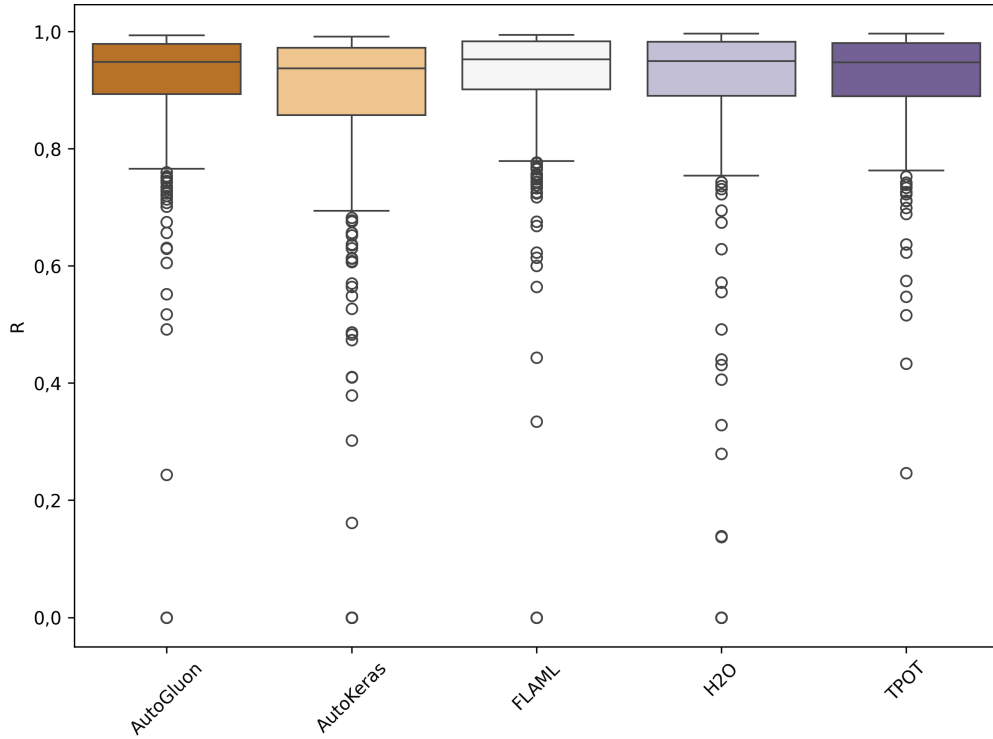


Fonte: Elaborado pelo autor (2025).

4.1 MÉTRICAS DE DESEMPENHO

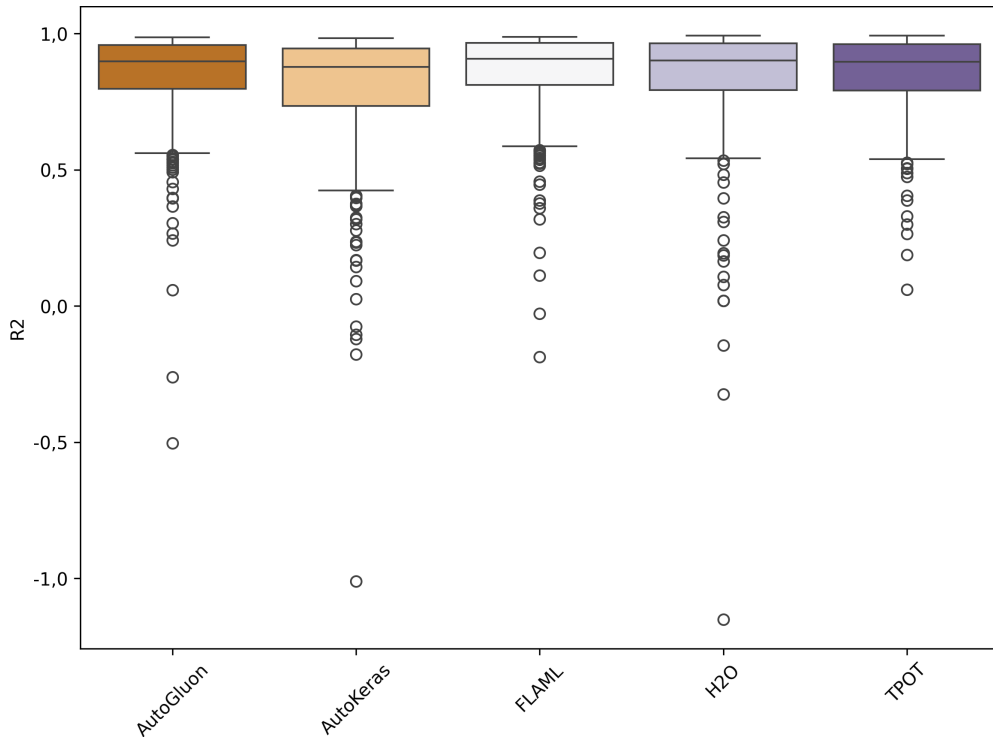
As Figuras 28 a 32 consolidam os resultados de todos os experimentos computacionais.

Figura 28 – Distribuição dos resultados por sistemas para a métrica R.



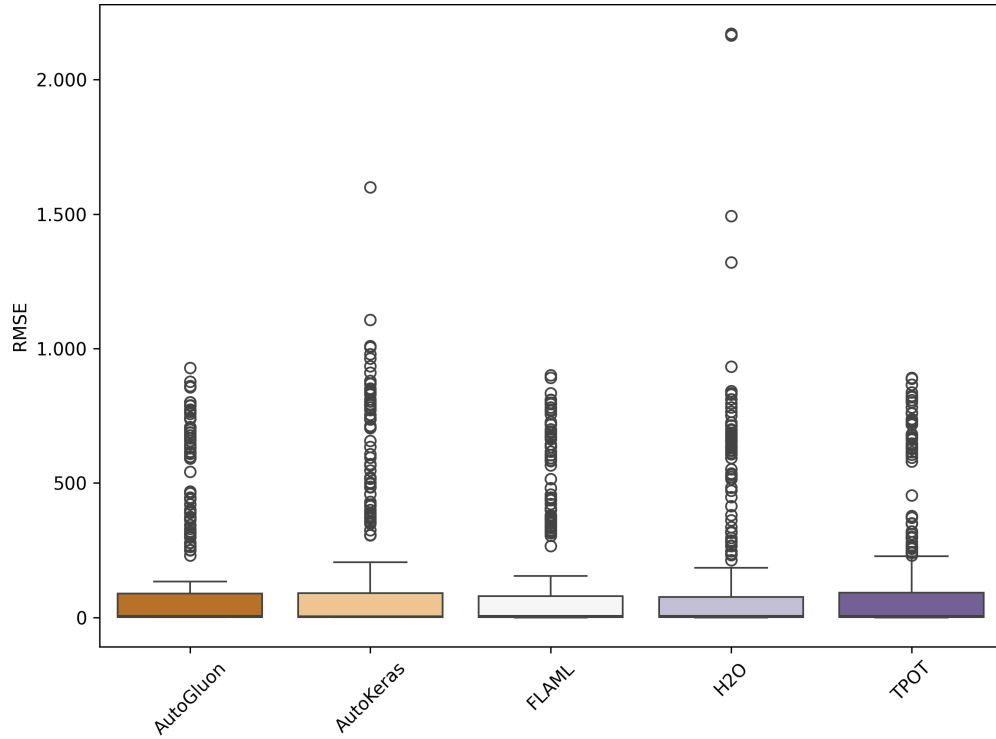
Fonte: Elaborado pelo autor (2026).

Figura 29 – Distribuição dos resultados por sistemas para a métrica R^2 .



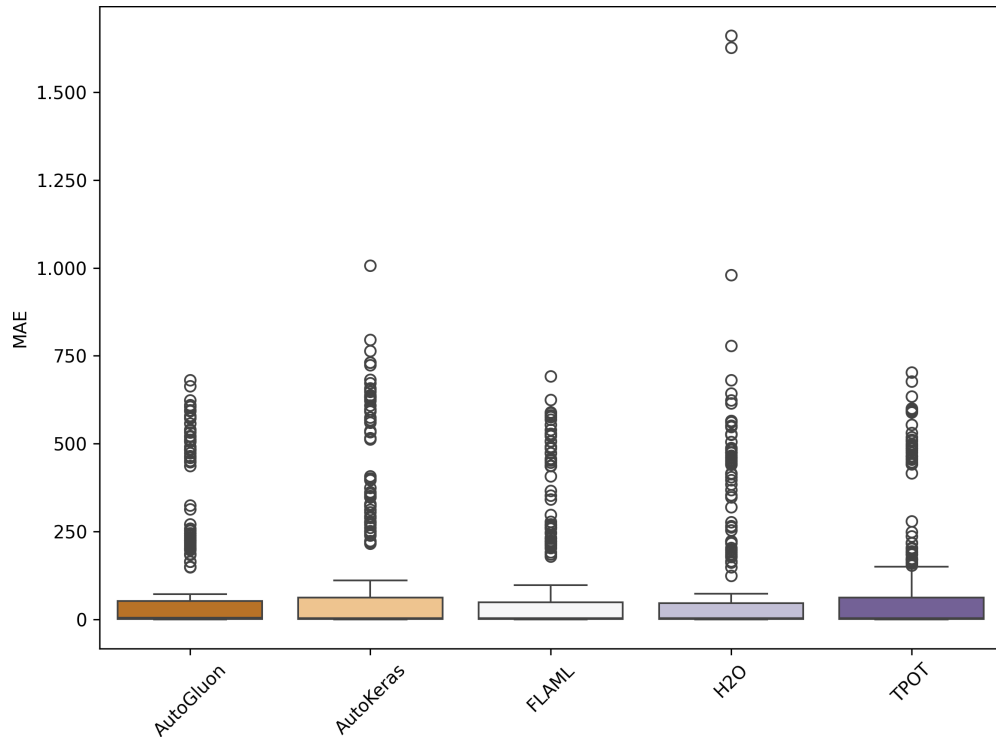
Fonte: Elaborado pelo autor (2026).

Figura 30 – Distribuição dos resultados por sistemas para a métrica RMSE.



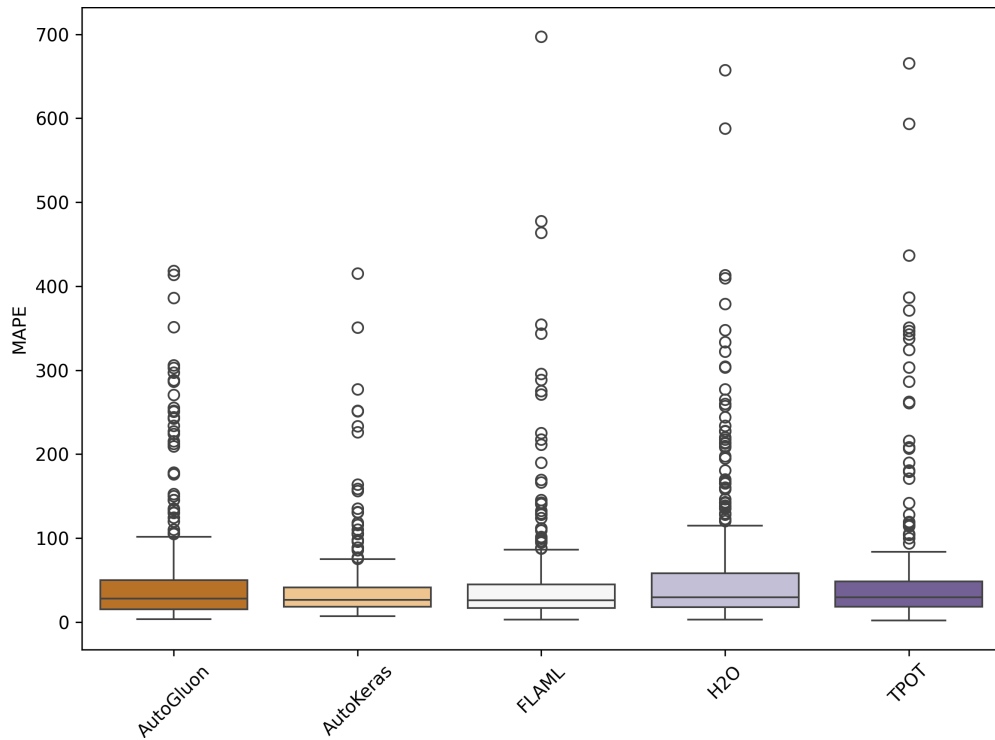
Fonte: Elaborado pelo autor (2026).

Figura 31 – Distribuição dos resultados por sistemas para a métrica MAE.



Fonte: Elaborado pelo autor (2026).

Figura 32 – Distribuição dos resultados por sistemas para a métrica MAPE.



Fonte: Elaborado pelo autor (2026).

Observa-se que, quando se avaliam todas as métricas, o modelo FLAML é o que possui menor variabilidade de resultados. Em contrapartida, o sistema AutoKeras foi o que apresentou maior variabilidade, com exceção apenas na métrica MAPE.

As Tabelas 17 e 18 reúnem os valores médios e as Tabelas 19 e 20 os DPs das métricas de desempenho obtidas em todos os conjuntos de dados avaliados, considerando os diferentes sistemas AutoML. Para fins de contextualização, também são apresentados os resultados dos modelos reportados nos estudos originais que deram origem aos conjuntos de dados, identificados como "Ref.". Esses valores correspondem, em geral, ao melhor desempenho alcançado nos respectivos trabalhos, cujos modelos estão descritos na Figura 33. Uma exceção ocorre no estudo de Wang and Yin (20), no qual os resultados publicados representam médias obtidas a partir de agrupamentos experimentais, e não o melhor desempenho individual. Adicionalmente, como Goliatt et al. (58) realizaram experimentos utilizando os conjuntos D1 a D6, seus resultados foram incluídos para ampliar a base comparativa da análise.

Tabela 17 – Valores médios das métricas de desempenho. (-) indica valores não disponíveis.

Conjunto de dados	Sistema/ Modelo	R	R ²	RMSE	MAE	MAPE (%)
D1	AutoGluon	0,904	0,819	0,234 MPa	0,163 MPa	128,371
	AutoKeras	0,902	0,816	0,234 MPa	0,152 MPa	92,089
	FLAML	0,912	0,835	0,217 MPa	0,146 MPa	137,759
	H2O	0,928	0,863	0,202 MPa	0,140 MPa	94,955
	TPOT	0,894	0,801	0,243 MPa	0,169 MPa	149,595
Ref. (48)	SVR	–	0,75	0,50 MPa	0,44 MPa	–
Goliatt et al.	XGB	0,964	0,928	0,148 MPa	0,101 MPa	–
D2	AutoGluon	0,752	0,525	0,875 MPa	0,617 MPa	42,235
	AutoKeras	0,867	0,744	0,584 MPa	0,403 MPa	32,049
	FLAML	0,778	0,579	0,820 MPa	0,582 MPa	40,229
	H2O	0,713	0,462	0,895 MPa	0,674 MPa	43,297
	TPOT	0,815	0,696	0,676 MPa	0,489 MPa	35,099
Ref. (48)	MLR	–	0,82	0,53 MPa	0,45 MPa	–
Goliatt et al.	XGB	0,910	0,826	0,596 MPa	0,390 MPa	–
D3	AutoGluon	0,901	0,814	0,489 MPa	0,376 MPa	25,594
	AutoKeras	0,889	0,763	0,538 MPa	0,375 MPa	21,883
	FLAML	0,908	0,825	0,475 MPa	0,367 MPa	24,596
	H2O	0,915	0,839	0,452 MPa	0,347 MPa	23,585
	TPOT	0,898	0,810	0,484 MPa	0,366 MPa	24,747
Ref. (49)	ANN	0,925	–	0,419 MPa	0,292 MPa	–
Goliatt et al.	XGB	0,940	0,882	0,410 MPa	0,312 MPa	–
D4	AutoGluon	0,967	0,936	19,316 kPa	14,137 kPa	5,747
	AutoKeras	0,685	0,456	55,800 kPa	43,646 kPa	17,903
	FLAML	0,980	0,961	14,758 kPa	11,042 kPa	4,597
	H2O	0,984	0,968	13,593 kPa	10,119 kPa	4,100
	TPOT	0,977	0,955	15,575 kPa	11,666 kPa	4,759
Ref. (50)	ANN	0,986	0,971	7,165 kPa	–	–
Goliatt et al.	XGB	0,994	0,987	7,005 kPa	6,527 kPa	–
D5	AutoGluon	0,981	0,963	1,233 MPa	0,809 MPa	182,095
	AutoKeras	0,984	0,968	1,140 MPa	0,697 MPa	79,015
	FLAML	0,984	0,969	1,132 MPa	0,703 MPa	119,320
	H2O	0,982	0,965	1,193 MPa	0,810 MPa	209,786
	TPOT	0,979	0,959	1,293 MPa	0,829 MPa	155,239
Ref. (51)	ANN	0,982	–	1225 MPa	–	83,4
Goliatt et al.	XGB	0,989	0,979	0,939 MPa	0,593 MPa	–

Fonte: Elaborado pelo autor (2025).

Tabela 18 – Valores médios das métricas de desempenho. (-) indica valores não disponíveis (continuação).

Conjunto de dados	Sistema/Modelo	R	R ²	RMSE	MAE	MAPE (%)
D6	AutoGluon	0,778	0,608	721,829 kPa	534,059 kPa	64,332
	AutoKeras	0,685	0,430	855,464 kPa	635,403 kPa	59,773
	FLAML	0,787	0,623	704,867 kPa	502,085 kPa	52,765
	H2O	0,797	0,638	692,648 kPa	496,498 kPa	58,102
	TPOT	0,780	0,611	720,204 kPa	522,665 kPa	56,909
Ref. (3)	BRT	–	0,69	649,73 kPa	472,33 kPa	–
Goliatt et al.	XGB	0,851	0,723	609,347 kPa	423,656 kPa	–
D7	AutoGluon	0,979	0,958	85,115 kPa	50,835 kPa	14,191
	AutoKeras	0,967	0,935	103,715 kPa	65,280 kPa	20,141
	FLAML	0,980	0,960	80,079 kPa	50,364 kPa	13,789
	H2O	0,969	0,940	93,932 kPa	59,752 kPa	19,262
	TPOT	0,969	0,940	99,597 kPa	65,090 kPa	20,110
Ref. (52)	SVR	0,997	–	–	–	4,606
D8	AutoGluon	0,953	0,908	15,200 MPa	10,428 MPa	14,894
	AutoKeras	0,950	0,902	15,282 MPa	11,458 MPa	14,614
	FLAML	0,943	0,890	16,671 MPa	11,721 MPa	15,474
	H2O	0,929	0,866	17,549 MPa	13,087 MPa	18,876
	TPOT	0,946	0,896	16,208 MPa	11,659 MPa	16,433
Ref. (53)	GPR	–	0,996	0,522 MPa	0,04 MPa	0,032
D9	AutoGluon	0,939	0,883	0,636 MPa	0,375 MPa	30,460
	AutoKeras	0,936	0,877	0,632 MPa	0,374 MPa	33,285
	FLAML	0,943	0,890	0,610 MPa	0,344 MPa	30,458
	H2O	0,928	0,862	0,677 MPa	0,415 MPa	38,987
	TPOT	0,924	0,854	0,705 MPa	0,441 MPa	41,195
Ref. (20)	MEP	–	0,857	0,652 MPa	0,415 MPa	–
D10	AutoGluon	0,986	0,971	364,036 kPa	220,969 kPa	26,285
	AutoKeras	0,977	0,954	451,564 kPa	299,045 kPa	36,417
	FLAML	0,984	0,968	380,999 kPa	237,462 kPa	21,450
	H2O	0,930	0,884	599,435 kPa	415,813 kPa	102,564
	TPOT	0,992	0,984	270,799 kPa	180,965 kPa	27,138
Ref. (9)	R \bar{F}	0,932	–	841,40 kPa	591,64 kPa	76,5

Fonte: Elaborado pelo autor (2025).

Tabela 19 – Desvios padrão das métricas de desempenho.

Conjunto de dados	Sistema/Modelo	R	R ²	RMSE	MAE	MAPE (%)
D1	AutoGluon	0,037	0,066	0,055 MPa	0,034 MPa	139,790
	AutoKeras	0,042	0,073	0,048 MPa	0,027 MPa	109,954
	FLAML	0,056	0,098	0,063 MPa	0,031 MPa	177,942
	H2O	0,031	0,057	0,043 MPa	0,026 MPa	110,567
	TPOT	0,043	0,076	0,050 MPa	0,030 MPa	186,377
D2	AutoGluon	0,161	0,323	0,257 MPa	0,201 MPa	14,832
	AutoKeras	0,146	0,261	0,184 MPa	0,120 MPa	13,506
	FLAML	0,152	0,276	0,203 MPa	0,158 MPa	14,214
	H2O	0,259	0,471	0,407 MPa	0,344 MPa	18,931
	TPOT	0,180	0,242	0,188 MPa	0,126 MPa	12,928
D3	AutoGluon	0,051	0,086	0,115 MPa	0,084 MPa	7,692
	AutoKeras	0,057	0,192	0,191 MPa	0,101 MPa	4,757
	FLAML	0,030	0,054	0,066 MPa	0,047 MPa	4,581
	H2O	0,038	0,067	0,063 MPa	0,048 MPa	4,219
	TPOT	0,064	0,106	0,122 MPa	0,073 MPa	5,713
D4	AutoGluon	0,018	0,034	6,020 kPa	3,778 kPa	1,565
	AutoKeras	0,176	0,259	12,711 kPa	10,590 kPa	4,246
	FLAML	0,014	0,028	5,116 kPa	3,477 kPa	1,613
	H2O	0,008	0,016	3,364 kPa	2,307 kPa	0,934
	TPOT	0,016	0,031	5,661 kPa	4,100 kPa	1,726
D5	AutoGluon	0,007	0,014	0,228 MPa	0,130 MPa	68,100
	AutoKeras	0,007	0,013	0,205 MPa	0,124 MPa	52,257
	FLAML	0,006	0,011	0,198 MPa	0,107 MPa	58,293
	H2O	0,007	0,015	0,231 MPa	0,144 MPa	80,903
	TPOT	0,008	0,017	0,242 MPa	0,168 MPa	89,615
D6	AutoGluon	0,053	0,081	92,884 kPa	64,614 kPa	19,530
	AutoKeras	0,107	0,302	181,128 kPa	107,515 kPa	22,157
	FLAML	0,062	0,096	101,052 kPa	76,596 kPa	14,022
	H2O	0,051	0,080	91,517 kPa	66,359 kPa	17,830
	TPOT	0,048	0,075	91,464 kPa	74,257 kPa	23,060
D7	AutoGluon	0,011	0,021	25,128 kPa	11,381 kPa	2,241
	AutoKeras	0,019	0,036	36,951 kPa	18,708 kPa	8,482
	FLAML	0,012	0,024	23,704 kPa	13,255 kPa	2,689
	H2O	0,046	0,082	64,250 kPa	39,038 kPa	20,820
	TPOT	0,018	0,035	32,389 kPa	18,787 kPa	6,402

Fonte: Elaborado pelo autor (2026).

Tabela 20 – Desvios padrão das métricas de desempenho (continuação).

Conjunto de dados	Sistema/Modelo	R	R ²	RMSE	MAE	MAPE (%)
D8	AutoGluon	0,019	0,036	3,265 MPa	2,146 MPa	3,236
	AutoKeras	0,028	0,052	4,153 MPa	3,587 MPa	4,092
	FLAML	0,023	0,043	2,978 MPa	1,933 MPa	2,968
	H2O	0,059	0,104	6,856 MPa	5,566 MPa	8,460
	TPOT	0,021	0,039	3,219 MPa	2,386 MPa	3,760
D9	AutoGluon	0,021	0,039	0,167 MPa	0,063 MPa	4,790
	AutoKeras	0,026	0,048	0,091 MPa	0,047 MPa	8,954
	FLAML	0,022	0,040	0,155 MPa	0,054 MPa	5,510
	H2O	0,039	0,071	0,203 MPa	0,120 MPa	17,640
	TPOT	0,029	0,052	0,166 MPa	0,080 MPa	11,662
D10	AutoGluon	0,005	0,010	95,065 kPa	40,940 kPa	10,078
	AutoKeras	0,011	0,021	93,672 kPa	52,739 kPa	21,320
	FLAML	0,004	0,007	65,827 kPa	37,877 kPa	4,962
	H2O	0,144	0,205	519,115 kPa	385,779 kPa	152,366
	TPOT	0,003	0,005	65,340 kPa	35,406 kPa	12,898

Fonte: Elaborado pelo autor (2026).

4.1.1 ANÁLISE DESCRITIVA

No conjunto D1, todos os sistemas alcançaram valores elevados de R² e R, superiores a 80%. O H2O apresentou os menores erros absolutos, representados por RMSE e MAE, enquanto o AutoKeras obteve o menor MAPE. No conjunto D2, observou-se maior dispersão nos resultados. AutoKeras e TPOT apresentaram melhor desempenho na maioria das métricas, com vantagem global para o AutoKeras. FLAML e H2O, por outro lado, apresentaram desempenho inferior.

No conjunto D3, os valores de R² foram superiores a 81% para todos os sistemas, com exceção do AutoKeras, que obteve 76,3%. AutoKeras e AutoGluon apresentaram maiores erros em termos de RMSE e MAE. O H2O registrou os melhores valores de R², R, RMSE e MAE, enquanto o AutoKeras manteve vantagem apenas no MAPE. Para o conjunto D4, a maioria dos sistemas apresentou resultados estáveis, exceto o AutoKeras, que obteve R² igual a 45,6%. Ao desconsiderar o AutoKeras, o conjunto D4 apresentou o melhor desempenho geral entre todos os conjuntos analisados.

No conjunto D5, os resultados foram consistentes entre os sistemas, com exceção do AutoGluon e do H2O, que apresentaram valores elevados de MAPE, iguais a 182,1% e 209,8%, respectivamente. No conjunto D6, os desempenhos foram mais modestos. O AutoKeras apresentou os piores resultados, enquanto o H2O obteve o melhor desempenho geral. Ainda assim, o FLAML apresentou o menor MAPE.

Figura 33 – Melhores modelos utilizados.

Conjunto de dados	Referência	Modelo
D1	Gajurel et al.	Regressão de Vetor Suporte (SVR)
D2	Gajurel et al.	Regressão Linear Múltipla (MLR)
D3	Ngo et al.	Rede Neural Artificial (ANN)
D4	Priyadarshee et al.	Rede Neural Artificial (ANN)
D5	Mozumder and Laskar	Rede Neural Artificial (ANN)
D6	Taffese and Abegaz	Árvore de Regressão de Agregação (BRT)
D7	Tabarsa et al.	Regressão de Vetor Suporte (SVR)
D8	Mahmoodzadeh et al.	Regressão do Processo Gaussiano (GPR)
D9	Wang and Yin	Programação de Múltiplas Expressões (MEP)
D10	Zhang et al.	Floresta Aleatória (RF)

Fonte: Elaborado pelo autor (2025).

No conjunto D7, todos os sistemas apresentaram resultados satisfatórios, com destaque para o AutoGluon. No conjunto D8, apesar do bom desempenho geral, o TPOT superou os demais sistemas em todas as métricas.

No conjunto D9, AutoGluon e FLAML apresentaram os melhores resultados, com vantagem clara do FLAML em todas as métricas. H2O e TPOT apresentaram desempenho inferior. No conjunto D10, o H2O apresentou desempenho reduzido em relação aos demais, enquanto o TPOT obteve os melhores resultados na maioria das métricas. Para MAPE, o FLAML apresentou valores inferiores.

Considerando os conjuntos D5, D7, D8, D9 e D10, todos os sistemas apresentaram valores elevados de R^2 . O conjunto D5 destacou-se por apresentar os maiores valores de R^2 e R entre todos os conjuntos. Em contraste, os conjuntos D2 e D6 apresentaram os menores valores dessas métricas.

Ao restringir a análise aos conjuntos D1 a D6, o método apresentado por Goliatt et al. (58) apresentou os melhores resultados entre todas as abordagens consideradas. Em uma avaliação adicional, que engloba todos os conjuntos de dados e desconsidera este estudo específico, verifica-se que os sistemas AutoML, de forma consistente, alcançaram desempenho superior ao dos modelos originalmente reportados na literatura. Essa comparação é apresentada na Tabela 21.

Para essa análise, foi adotado um critério de pontuação binária para cada sistema AutoML. A avaliação foi realizada métrica a métrica e conjunto a conjunto. Em cada caso, o valor médio foi confrontado com o melhor resultado publicado no estudo de referência correspondente. Atribuiu-se o valor “1” quando o sistema AutoML superou o desempenho do modelo de referência. Caso contrário, foi atribuído o valor “0”. A linha denominada “AutoML” corresponde à soma das pontuações obtidas por todos os sistemas

empregados neste trabalho (AutoGluon, AutoKeras, FLAML, H2O e TPOT) para cada métrica avaliada. Por sua vez, a linha “Referência” indica o número de conjuntos de dados em que os estudos originais apresentaram os melhores resultados (os modelos de referência e seus respectivos estudos podem ser consultados na Figura 33). A coluna “Pontuação” sintetiza o desempenho global, ao contabilizar o total de métricas nas quais cada abordagem superou os modelos de referência.

Tabela 21 – Comparação binária de desempenho entre os sistemas AutoML e os modelos de referência.

Modelo/Sistema	R	R ²	RMSE	MAE	MAPE	Pontuação
AutoGluon	0	0	0	0	0	0
AutoKeras	0	1	0	1	1	3
FLAML	1	1	2	1	1	6
H2O	0	1	2	1	0	4
TPOT	1	0	1	1	0	3
AutoML	2	3	5	4	2	16
Referência	3	3	4	3	2	15

Fonte: Elaborado pelo autor (2025).

A Tabela 22 reúne os valores médios do ID, acompanhados de seus DPs. Em cada conjunto, o maior valor médio de ID foi destacado. De forma complementar, a Tabela 23 identifica, para cada conjunto de dados, o sistema que apresentou o melhor desempenho médio segundo esse índice.

Note que há diferenças claras no comportamento dos sistemas. O FLAML apresentou o maior ID médio global (0,7848), além de baixa dispersão dos resultados, com DPs médios próximos de 0,0809. Esses valores indicam estabilidade ao longo das execuções e eficiência no processo de seleção de modelos. O desempenho do FLAML foi particularmente consistente em bases com maior número de amostras, como D5, D6, D7 e D9, sugerindo melhor aproveitamento de conjuntos de dados mais extensos.

O AutoKeras apresentou maior sensibilidade às características dos conjuntos de dados analisados. O ID médio foi inferior (0,7156), com desempenho significativamente reduzido no conjunto D4 (0,413). Esse resultado pode estar associado à sua estratégia centrada na busca por arquiteturas de redes neurais, que nem sempre se adapta adequadamente a dados tabulares. Apesar do desempenho global inferior do AutoKeras, ele apresentou melhores resultados em comparação aos demais sistemas nos conjuntos D2 e D8. Vale ressaltar que o conjunto D8 se diferencia dos demais por ser o único que apresenta rochas.

Tabela 22 – Desempenho médio (ID) e desvio padrão dos sistemas AutoML para cada conjunto de dados.

Conjunto de dados	AutoGluon	AutoKeras	FLAML	H2O	TPOT
D1	0,486 (±0,148)	0,498 (±0,136)	0,536 (±0,182)	0,581 (±0,121)	0,451 (±0,146)
D2	0,610 (±0,082)	0,706 (±0,072)	0,625 (±0,076)	0,578 (±0,184)	0,676 (±0,067)
D3	0,797 (±0,111)	0,787 (±0,156)	0,808 (±0,062)	0,829 (±0,064)	0,801 (±0,106)
D4	0,892 (±0,049)	0,413 (±0,197)	0,931 (±0,046)	0,944 (±0,028)	0,923 (±0,054)
D5	0,667 (±0,143)	0,748 (±0,134)	0,748 (±0,117)	0,685 (±0,151)	0,625 (±0,173)
D6	0,764 (±0,069)	0,676 (±0,150)	0,802 (±0,074)	0,800 (±0,076)	0,780 (±0,079)
D7	0,894 (±0,049)	0,831 (±0,079)	0,897 (±0,056)	0,847 (±0,208)	0,833 (±0,088)
D8	0,788 (±0,089)	0,792 (±0,112)	0,741 (±0,092)	0,687 (±0,243)	0,746 (±0,100)
D9	0,786 (±0,093)	0,778 (±0,082)	0,810 (±0,091)	0,716 (±0,199)	0,684 (±0,128)
D10	0,954 (±0,017)	0,927 (±0,022)	0,950 (±0,013)	0,850 (±0,223)	0,972 (±0,014)

Fonte: Elaborado pelo autor (2025).

Tabela 23 – Resumo dos melhores sistemas AutoML por conjunto de dados.

Conjunto de dados	Melhor sistema	ID médio
D1	H2O	0,581
D2	AutoKeras	0,706
D3	H2O	0,829
D4	H2O	0,944
D5	FLAML	0,748
D6	FLAML	0,802
D7	FLAML	0,897
D8	AutoKeras	0,792
D9	FLAML	0,810
D10	TPOT	0,972

Fonte: Elaborado pelo autor (2025).

Os sistemas AutoGluon, H2O e TPOT apresentaram desempenhos intermediários, com valores médios de ID de 0,7638, 0,7517 e 0,7491, respectivamente, indicando resultados competitivos, porém menos uniformes.

4.1.2 ANÁLISE ESTATÍSTICA

Para cada conjunto de dados e métrica, foram avaliadas a normalidade e a homogeneidade das variâncias por meio dos testes de Shapiro–Wilk e Levene. Os resultados completos são apresentados no Apêndice A. Em alguns casos pontuais, como algumas métricas do conjunto D1, essas condições foram atendidas para determinados sistemas. No entanto, testes paramétricos exigem que todas as amostras comparadas atendam simultaneamente a essas condições. Em nenhum dos cenários avaliados esse requisito foi plenamente satisfeito. Dessa forma, a aplicação do teste de análise de variância (ANOVA) não foi considerada adequada.

Em razão dessas limitações, optou-se pela utilização do teste de Kruskal-Wallis, que não depende das hipóteses de normalidade nem de homogeneidade das variâncias. Esse teste forneceu uma base consistente para a comparação do desempenho preditivo entre os sistemas AutoML.

O teste de Kruskal-Wallis foi aplicado em todas as métricas e conjuntos de dados. A Figura 34 resume os casos em que foram identificadas diferenças estatisticamente significativas ($p < 0,05$). Observou-se que os conjuntos D2, D4, D5, D7 e D10 apresentaram diferenças significativas em todas as métricas avaliadas. Em contraste, os conjuntos D3 e D8 não apresentaram diferenças significativas em nenhuma métrica.

Figura 34 – Diferenças significativas do teste de Kruskal-Wallis entre sistemas.

Métrica	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
R	Sim	Sim	Não	Sim	Sim	Sim	Sim	Não	Sim	Sim
R ²	Sim	Sim	Não	Sim	Sim	Sim	Sim	Não	Sim	Sim
RMSE	Sim	Sim	Não	Sim	Sim	Sim	Sim	Não	Não	Sim
MAE	Sim	Sim	Não	Sim	Sim	Sim	Sim	Não	Sim	Sim
MAPE	Não	Sim	Não	Sim	Sim	Não	Sim	Não	Sim	Sim

Fonte: Elaborado pelo autor (2025).

Nos casos em que o teste de Kruskal-Wallis indicou diferenças significativas, foi aplicado o teste *post-hoc* de Dunn com correção de Bonferroni. Os resultados completos são apresentados na Figura 35, nos quais a direção da diferença mostra que os modelos à esquerda possuem melhores resultados, pois para a métrica R e R² são desejáveis maiores valores, enquanto o oposto acontece para as demais métricas. Nos casos em que o teste de Dunn não apresentou diferenças significativas entre alguns sistemas, mas que apresentou no teste de Kruskal-Wallis, são descritos a seguir: conjunto D5 para as métricas R e R²;

conjunto D7 para a métrica RMSE; e o conjunto D2 para a métrica MAPE. Os nomes dos sistemas foram abreviados da seguinte forma: AG = AutoGluon; AK = AutoKeras; FL = FLAML; H2 = H2O; TP = TPOT.

Figura 35 – Diferenças significativas entre pares do teste de Dunn entre sistemas.

Métrica	Conjunto	Direção da diferença
R	D1	H2 > TP
R	D2	AK > AG; AK > FL; AK > H2
R	D4	AG > AK; FL > AG; H2 > AG; FL > AK; H2 > AK; TP > AK
R	D6	AG > AK; FL > AK; H2 > AK; TP > AK
R	D7	FL > AK
R	D9	FL > TP
R	D10	AG > AK; TP > AG; TP > AK; TP > FL; TP > H2
R ²	D1	H2 > TP
R ²	D2	AK > AG; AK > FL; AK > H2
R ²	D4	AG > AK; FL > AG; H2 > AG; FL > AK; H2 > AK; TP > AK
R ²	D6	AG > AK; FL > AK; H2 > AK; TP > AK
R ²	D7	FL > AK
R ²	D9	FL > TP
R ²	D10	AG > AK; TP > AG; TP > AK; TP > FL; TP > H2
RMSE	D1	H2 < TP
RMSE	D2	AK < AG; TP < AG; AK < FL; AK < H2
RMSE	D4	AG < AK; FL < AG; H2 < AG; FL < AK; H2 < AK; TP < AK
RMSE	D5	AK < TP; FL < TP
RMSE	D6	AG < AK; FL < AK; H2 < AK; TP < AK
RMSE	D10	AG < AK; TP < AG; TP < AK; TP < FL; TP < H2
MAE	D1	FL < TP; H2 < TP
MAE	D2	AK < AG; AK < FL; AK < H2
MAE	D4	AG < AK; FL < AG; H2 < AG; FL < AK; H2 < AK; TP < AK
MAE	D5	AK < AG; FL < AG; AK < H2; AK < TP; FL < H2; FL < TP
MAE	D6	AG < AK; FL < AK; H2 < AK; TP < AK
MAE	D7	AG < AK; AG < TP; FL < AK; H2 < AK; FL < TP; H2 < TP
MAE	D9	AG < TP; AK < TP; FL < H2; FL < TP
MAE	D10	AG < AK; TP < AG; FL < AK; TP < AK; TP < FL; TP < H2
MAPE	D4	AG < AK; H2 < AG; FL < AK; H2 < AK; TP < AK
MAPE	D5	AK < AG; FL < AG; AK < H2; AK < TP; FL < H2
MAPE	D7	AG < AK; AG < TP; FL < AK; H2 < AK; FL < TP; H2 < TP
MAPE	D9	AG < TP; AK < TP; FL < TP
MAPE	D10	AG < H2; FL < AK; FL < H2; TP < H2

Fonte: Elaborado pelo autor (2025).

Ao usar a Figura 35 como referência, e contabilizar a quantidade de diferenças significativas entre sistemas AutoML por métrica, obtêm-se os valores da Tabela 24; e ao contabilizar a quantidade de diferenças significativas entre sistemas AutoML por conjuntos de dados, obtêm-se os valores da Tabela 25.

Tabela 24 – Quantidade de diferenças significativas entre sistemas por métrica, conforme o teste de Dunn.

Métrica	Quantidade
R	21
R ²	21
RMSE	22
MAE	37
MAPE	23

Fonte: Elaborado pelo autor (2026).

Tabela 25 – Quantidade de diferenças significativas entre sistemas por conjuntos de dados, conforme o teste de Dunn.

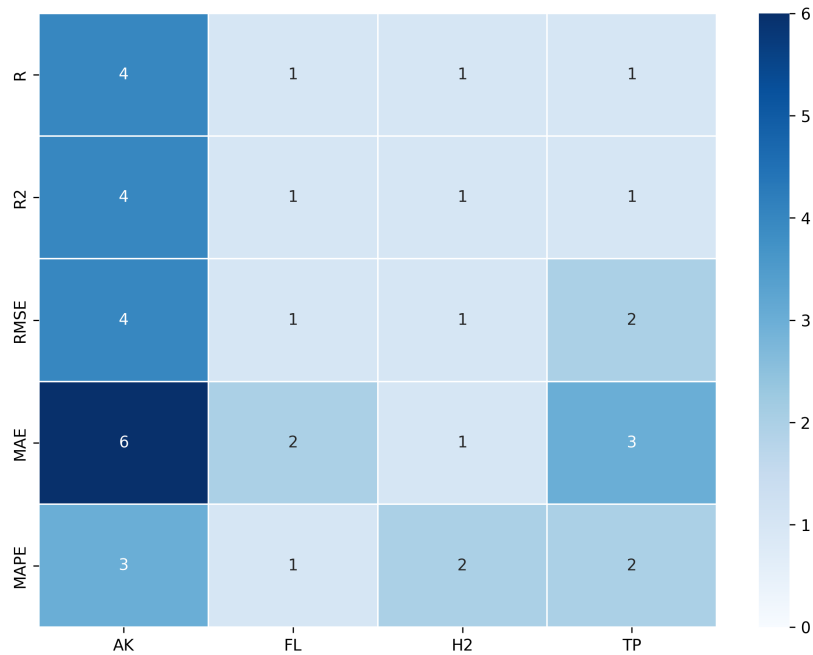
Conjunto de dados	Quantidade
D1	5
D2	13
D3	0
D4	29
D5	13
D6	16
D7	14
D8	0
D9	9
D10	25

Fonte: Elaborado pelo autor (2026).

As Figuras 36 a 40 sintetizam esses resultados por meio de mapas de calor. Cada figura corresponde a um sistema e indica o número de conjuntos de dados em que foram observadas diferenças estatisticamente significativas em relação aos demais sistemas. As linhas representam as métricas analisadas, enquanto as colunas indicam os sistemas comparados. Tonalidades mais escuras indicam maior número de comparações significativas.

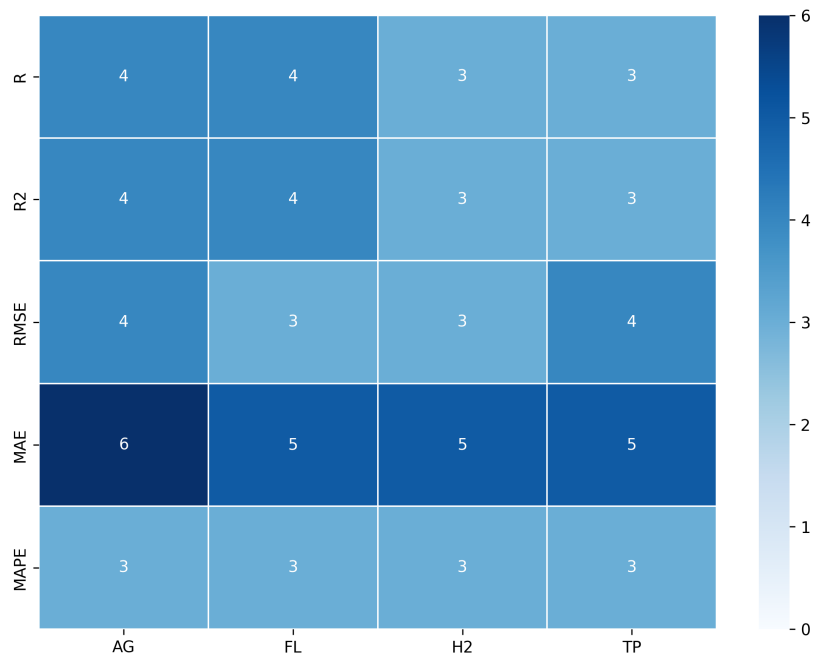
De forma geral, o AutoKeras apresentou diferenças significativas em relação a outros sistemas em diversos conjuntos de dados, em especial quando comparado ao AutoGluon. Os conjuntos D4 e D10 concentraram o maior número de diferenças estatísticas, indicando maior separação de desempenho entre os sistemas. Em contrapartida, os conjuntos D1, D3, D8 e em alguns casos o D5, apresentaram poucas diferenças, indicando desempenho mais próximo entre os sistemas.

Figura 36 – Mapa de calor da quantidade de diferenças significativas entre o AutoGluon e os demais sistemas.



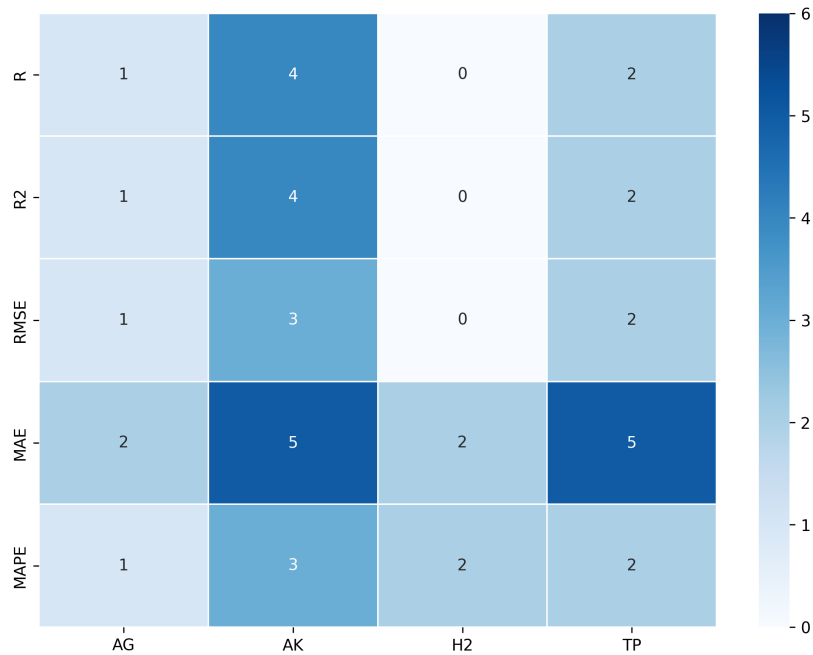
Fonte: Elaborado pelo autor (2025).

Figura 37 – Mapa de calor da quantidade de diferenças significativas entre o AutoKeras e os demais sistemas.



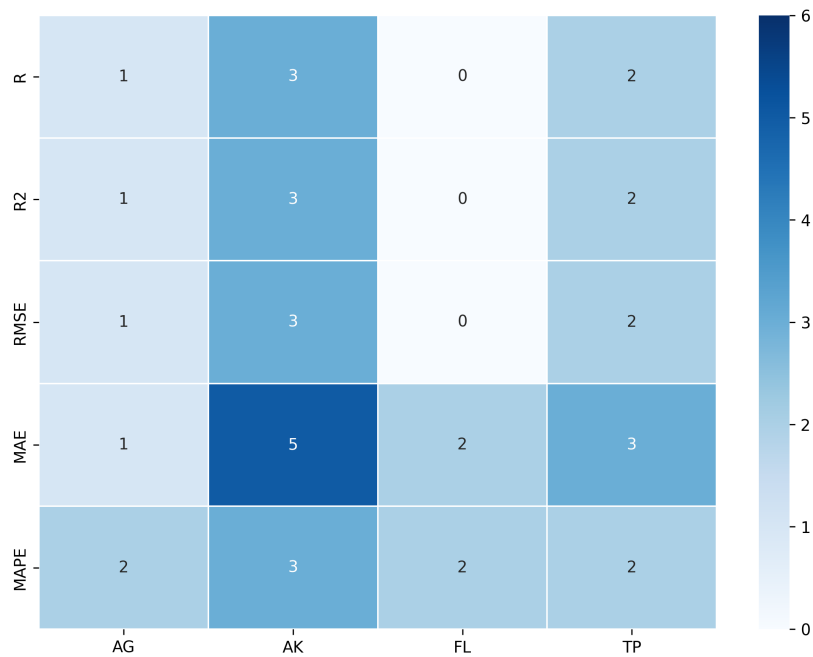
Fonte: Elaborado pelo autor (2025).

Figura 38 – Mapa de calor da quantidade de diferenças significativas entre o FLAML e os demais sistemas.



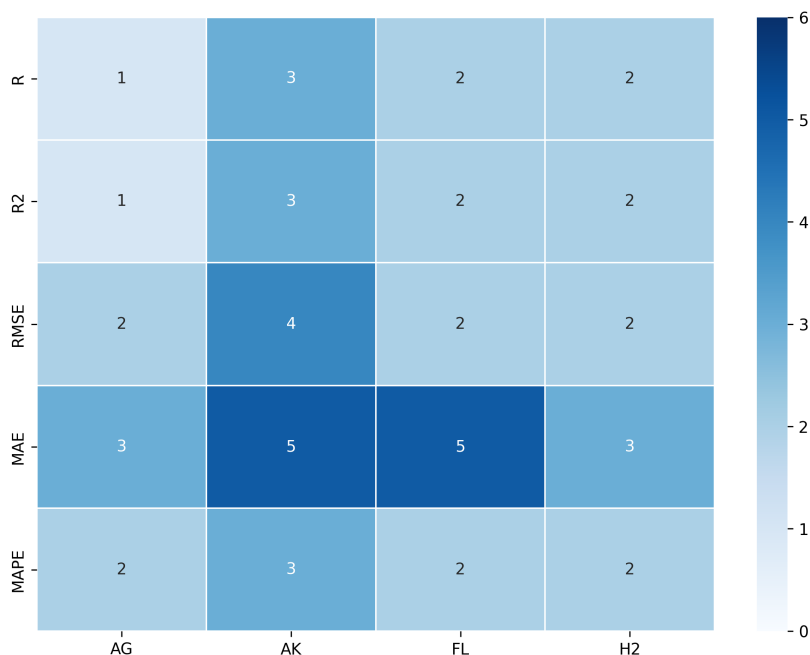
Fonte: Elaborado pelo autor (2025).

Figura 39 – Mapa de calor da quantidade de diferenças significativas entre o H2O e os demais sistemas.



Fonte: Elaborado pelo autor (2025).

Figura 40 – Mapa de calor da quantidade de diferenças significativas entre o TPOT e os demais sistemas.



Fonte: Elaborado pelo autor (2025).

4.2 IMPORTÂNCIA DOS ATRIBUTOS

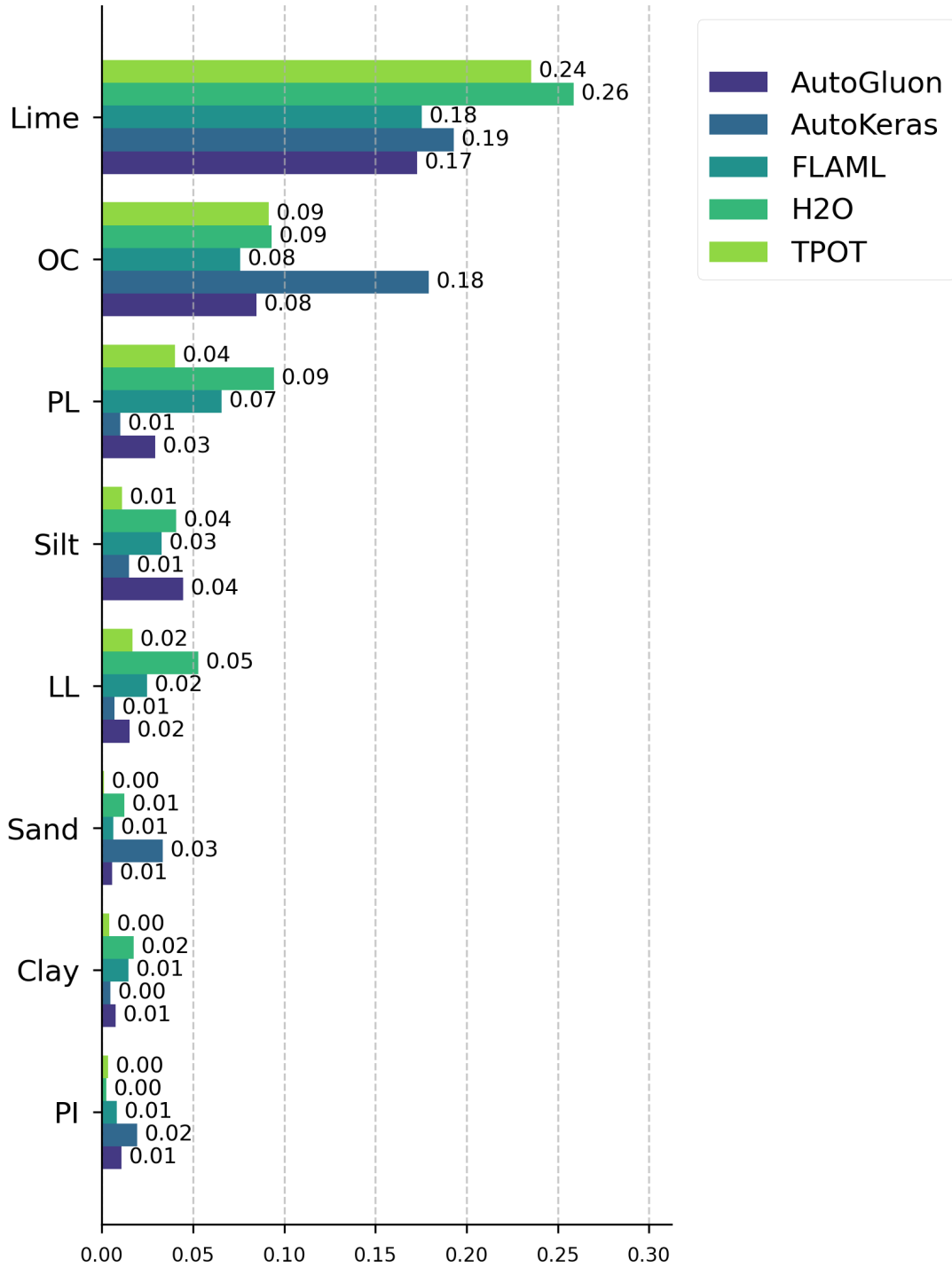
As Figuras 41 até 50 reúnem os valores médios de importância das variáveis de entrada para os conjuntos D1 a D10. A presença recorrente de variáveis ligadas ao teor de estabilizantes é mais frequente nos conjuntos D1, D2, D3, D4, D5 e D9. Já nos conjuntos D6 e D7, destacam-se variáveis associadas à compactação das amostras (peso específico seco, peso específico seco máximo e umidade ótima).

No conjunto D1, a seleção recorrente das variáveis "Lime", "PL" e "LL" sugere a obtenção de um bom desempenho do H2O. No conjunto D2, o TPOT priorizou de forma consistente o teor de cimento. No conjunto D3, o AutoKeras atribuiu elevada importância à variável relacionada à condição de cura. Situação semelhante ocorreu com o TPOT, que priorizou a variável de teor de GGBS no conjunto D5. Entretanto, situação oposta ocorre no conjunto D4, no qual o AutoKeras praticamente não selecionou a variável associada ao teor de casca de arroz, o que pode ter causado desempenho inferior em relação aos demais sistemas (Tabela 22).

No conjunto D7, observa-se maior coerência entre a frequência de seleção das variáveis e o desempenho final. O AutoGluon obteve bons resultados mesmo com menor ênfase no peso específico seco, enquanto o FLAML, que apresentou o maior ID, selecionou esse atributo com maior frequência. No conjunto D8, a baixa priorização da porosidade

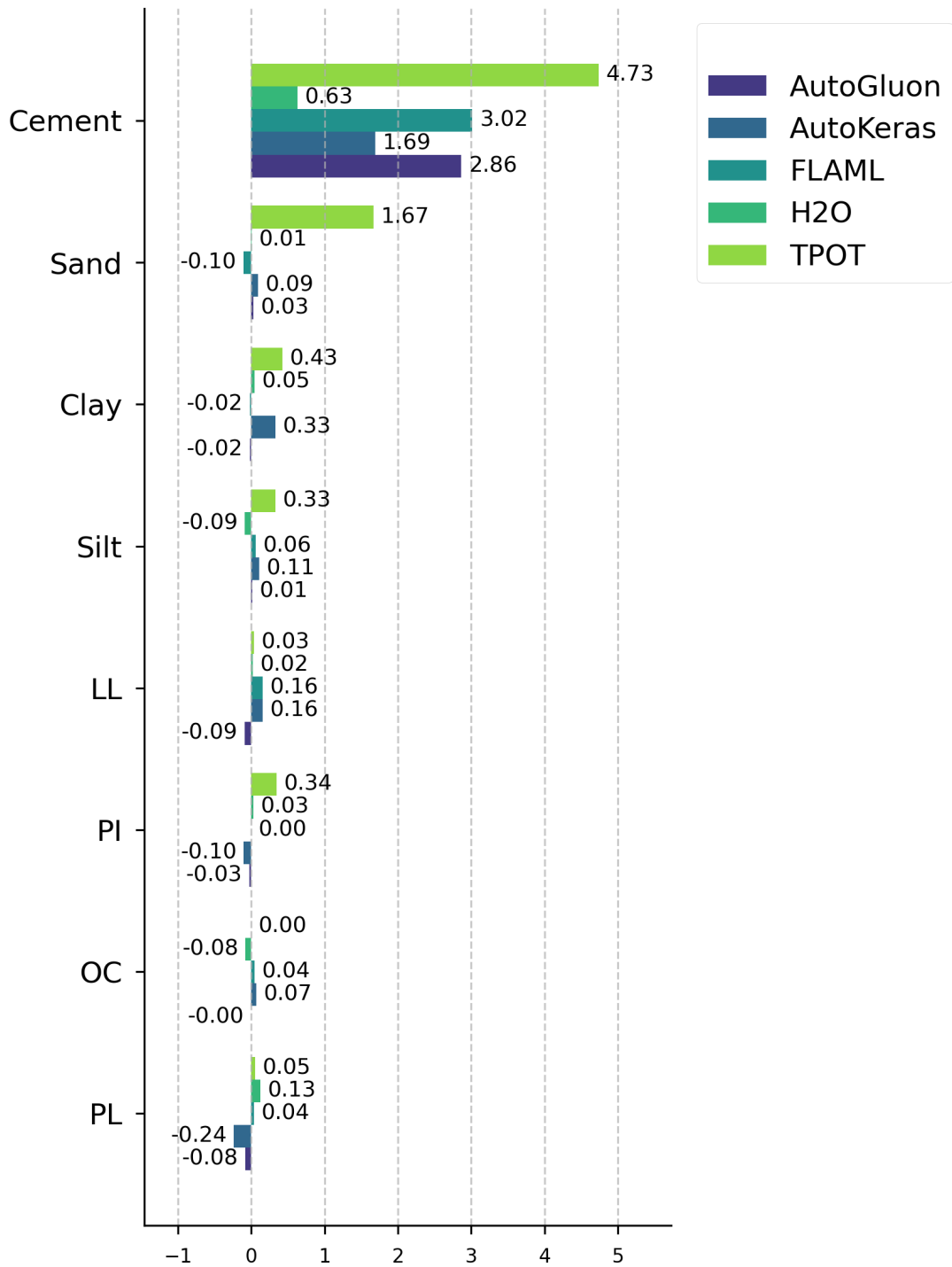
sugere uma contribuição para os resultados menos expressivos do H2O.

Figura 41 – Importância média dos atributos para o conjunto de dados D1.



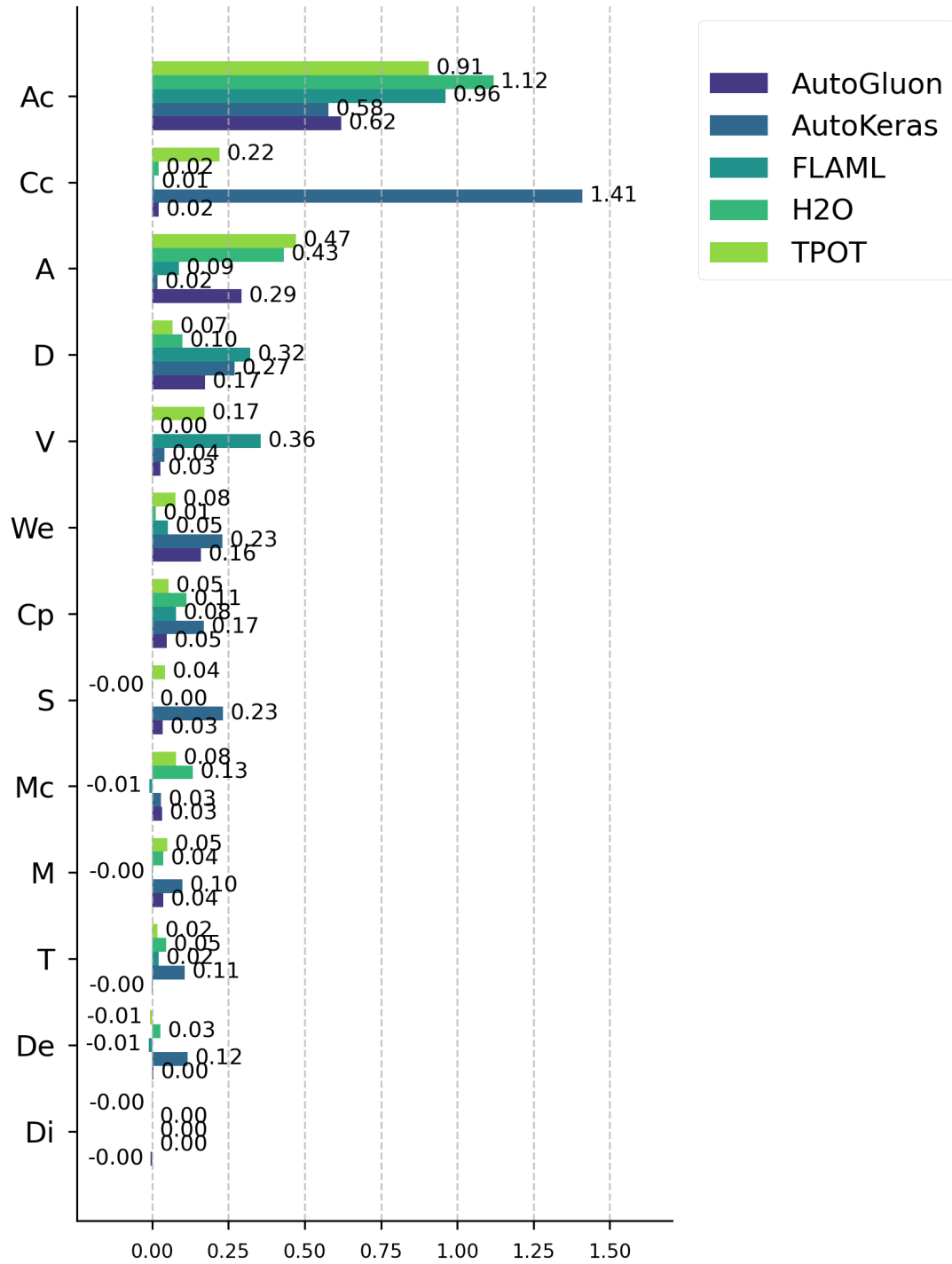
Fonte: Elaborado pelo autor (2025).

Figura 42 – Importância média dos atributos para o conjunto de dados D2.



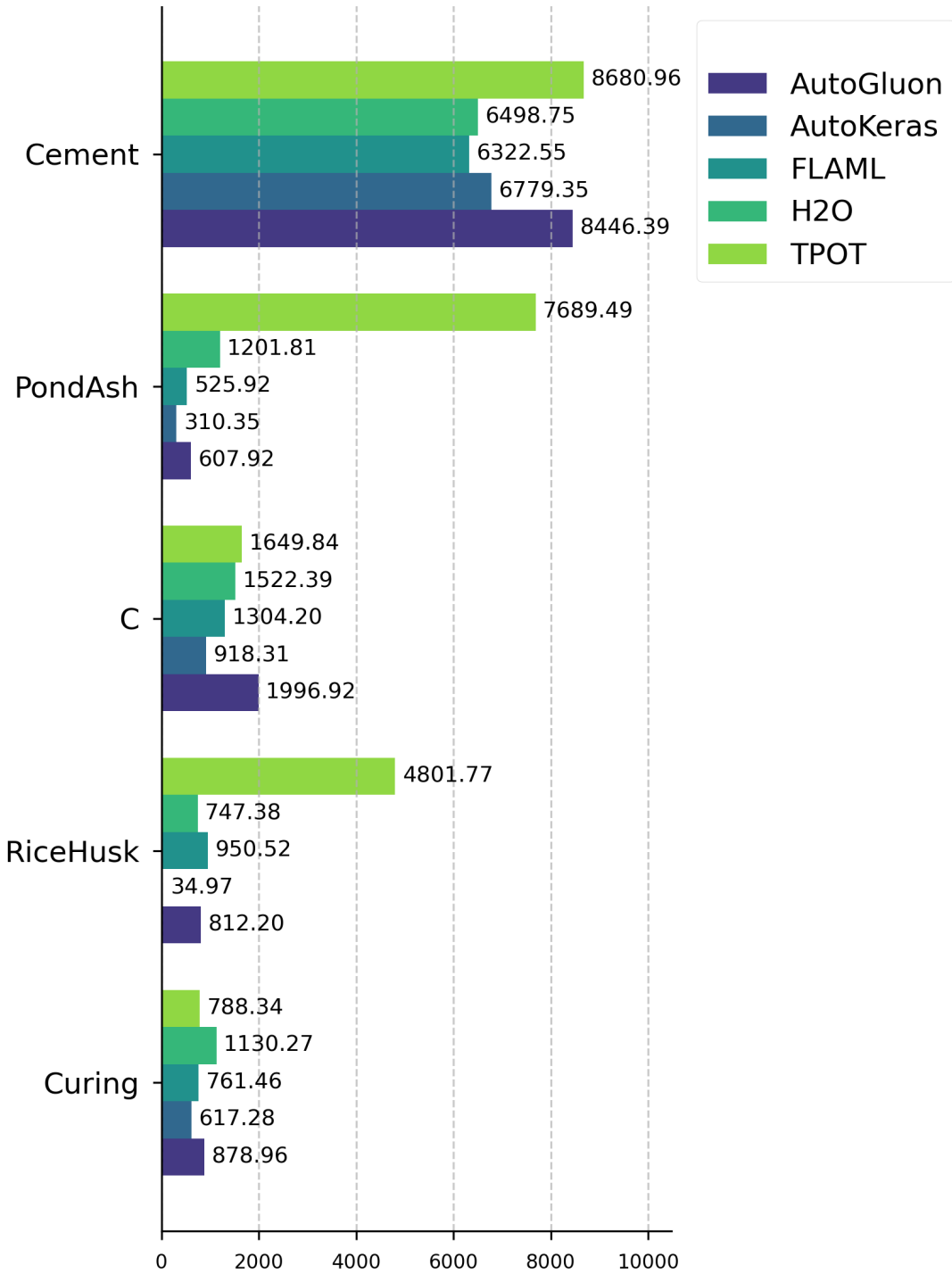
Fonte: Elaborado pelo autor (2025).

Figura 43 – Importância média dos atributos para o conjunto de dados D3.



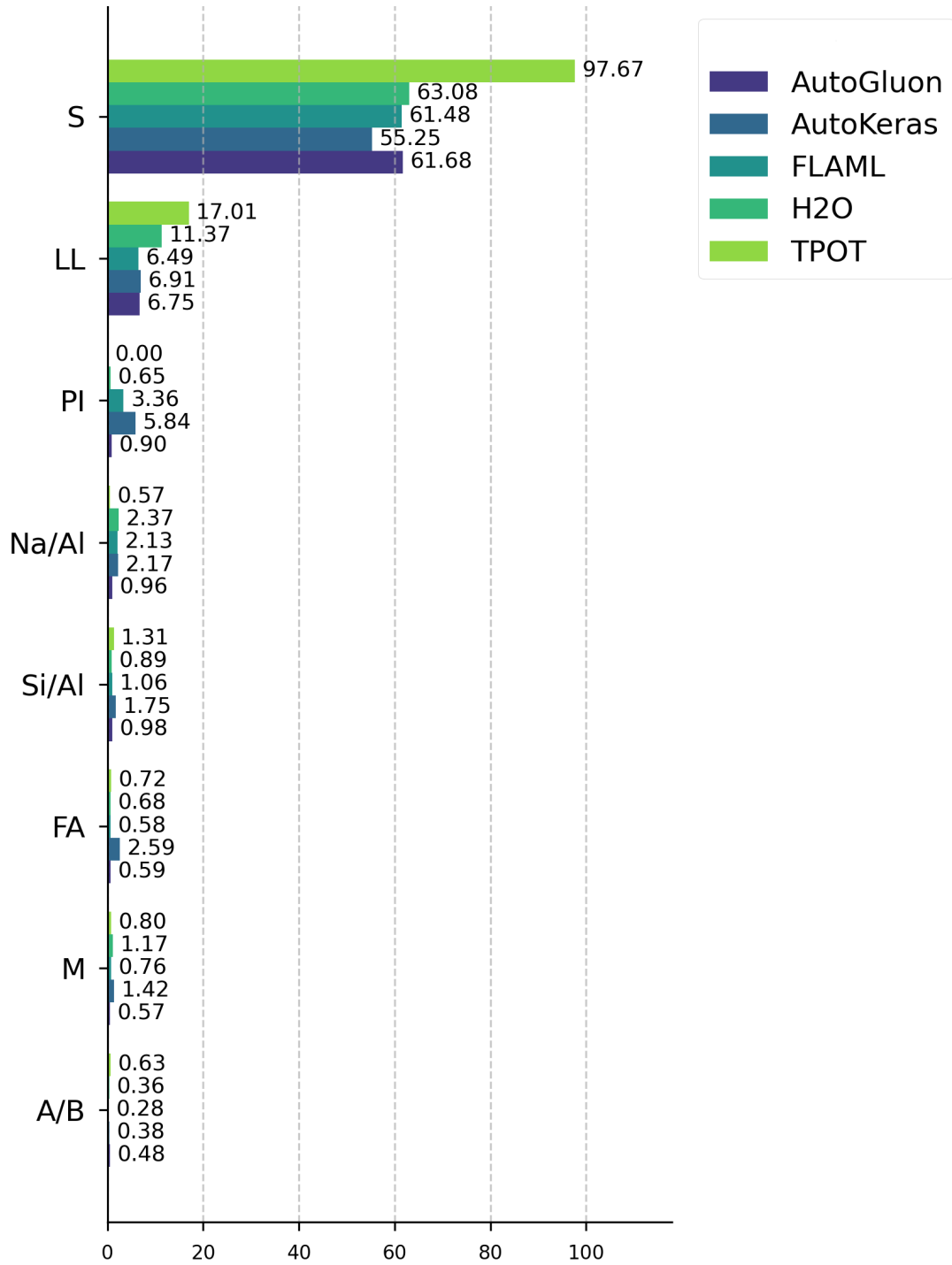
Fonte: Elaborado pelo autor (2025).

Figura 44 – Importância média dos atributos para o conjunto de dados D4.



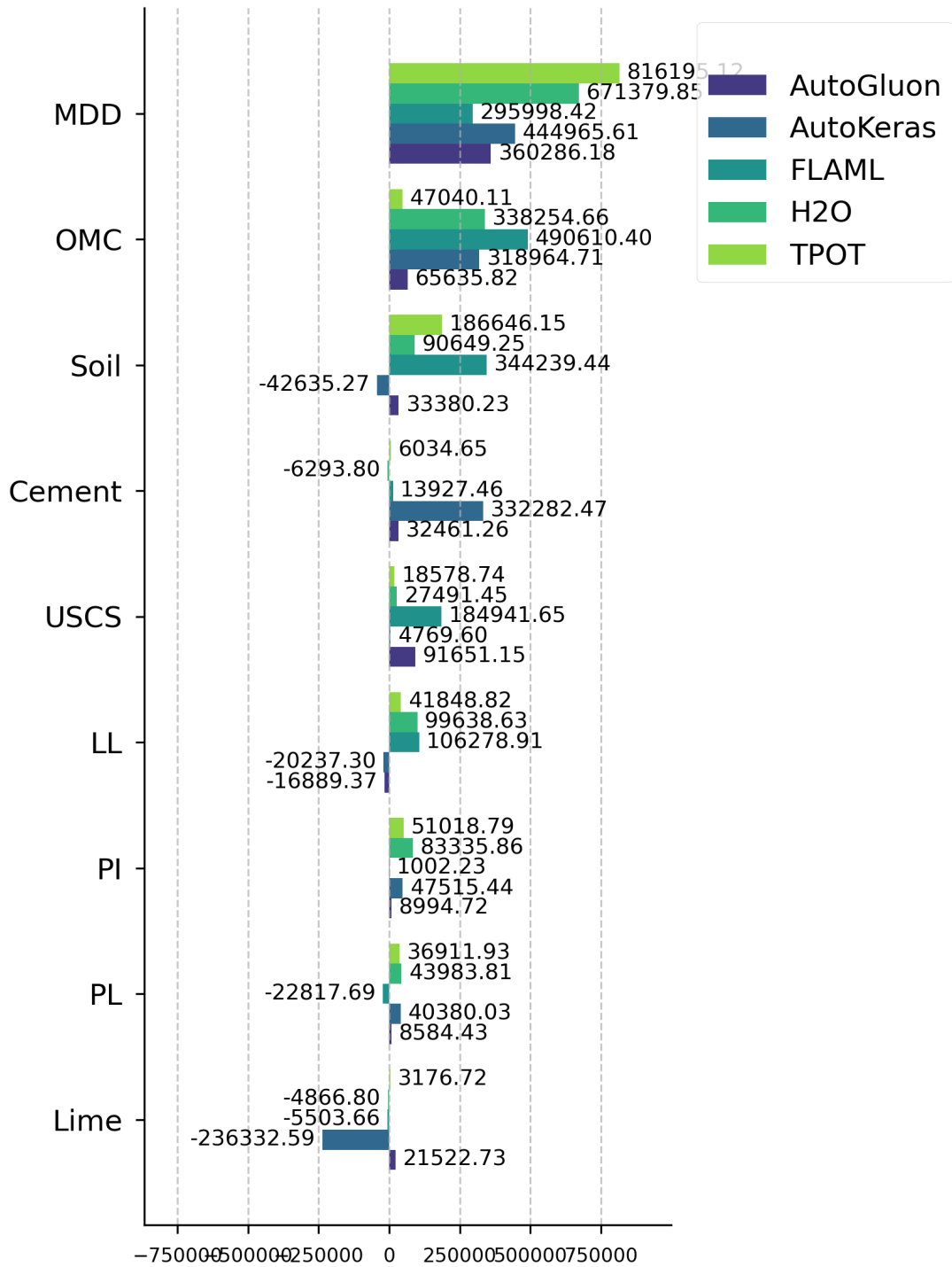
Fonte: Elaborado pelo autor (2025).

Figura 45 – Importância média dos atributos para o conjunto de dados D5.



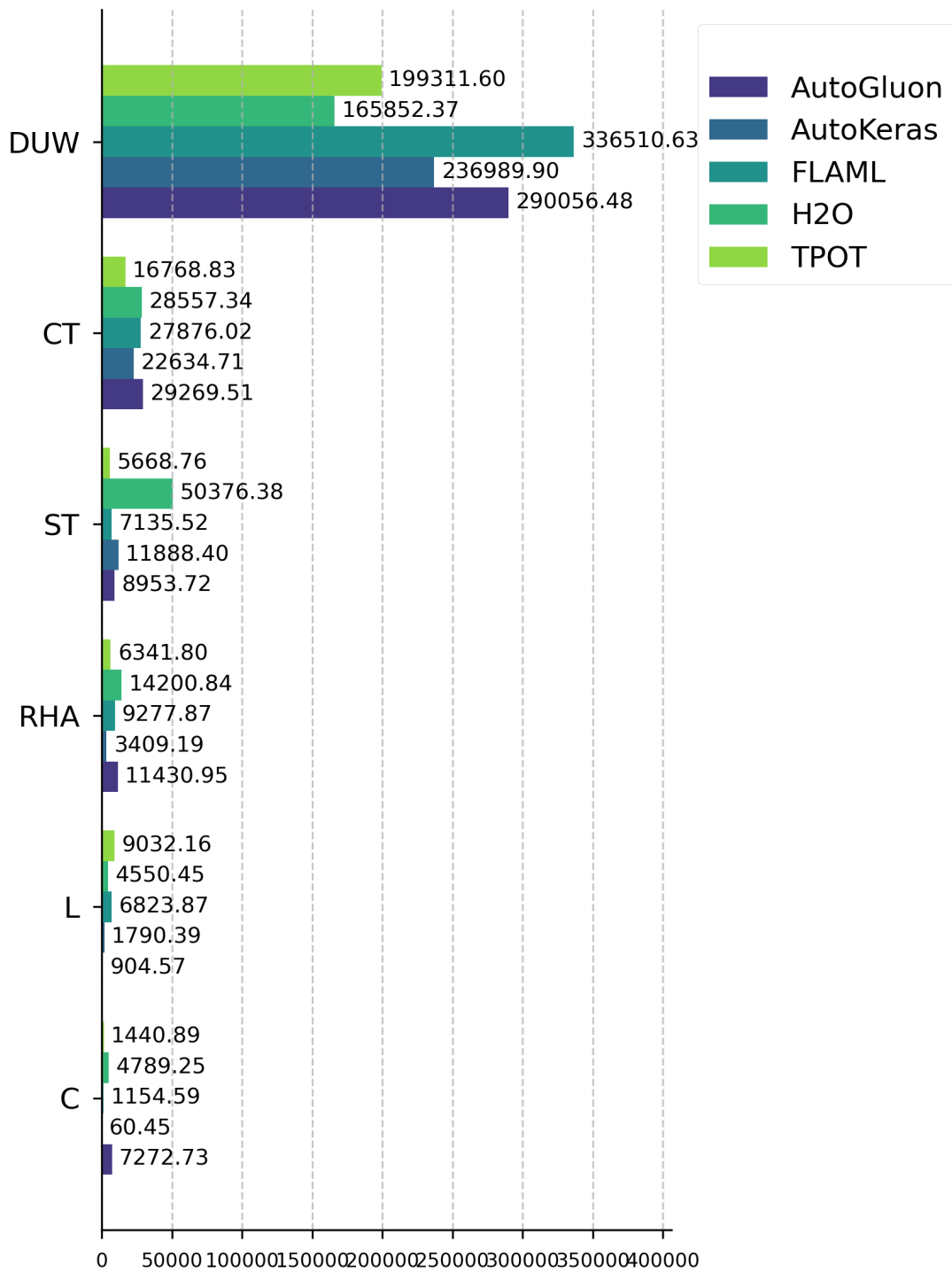
Fonte: Elaborado pelo autor (2025).

Figura 46 – Importância média dos atributos para o conjunto de dados D6.



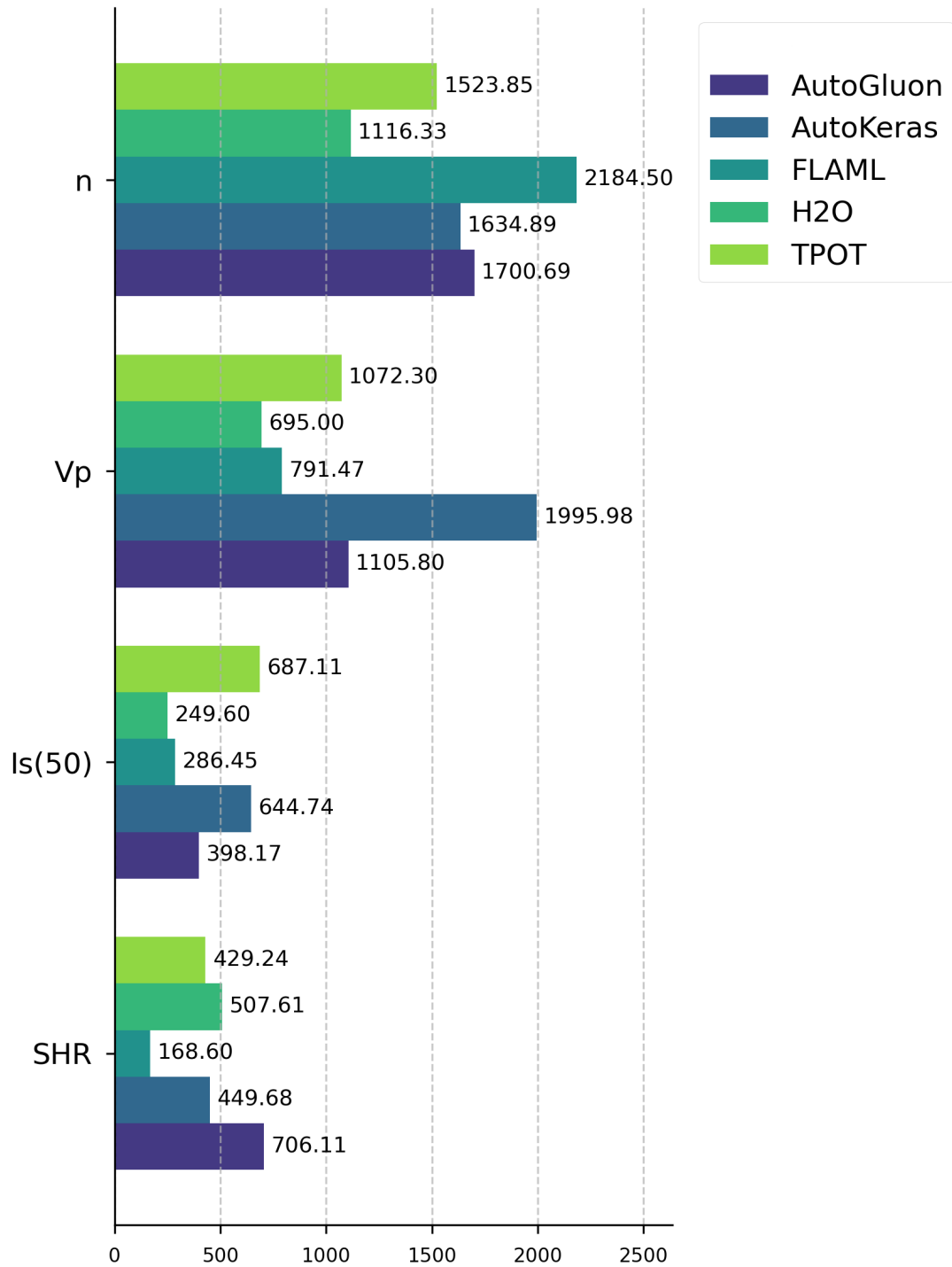
Fonte: Elaborado pelo autor (2025).

Figura 47 – Importância média dos atributos para o conjunto de dados D7.



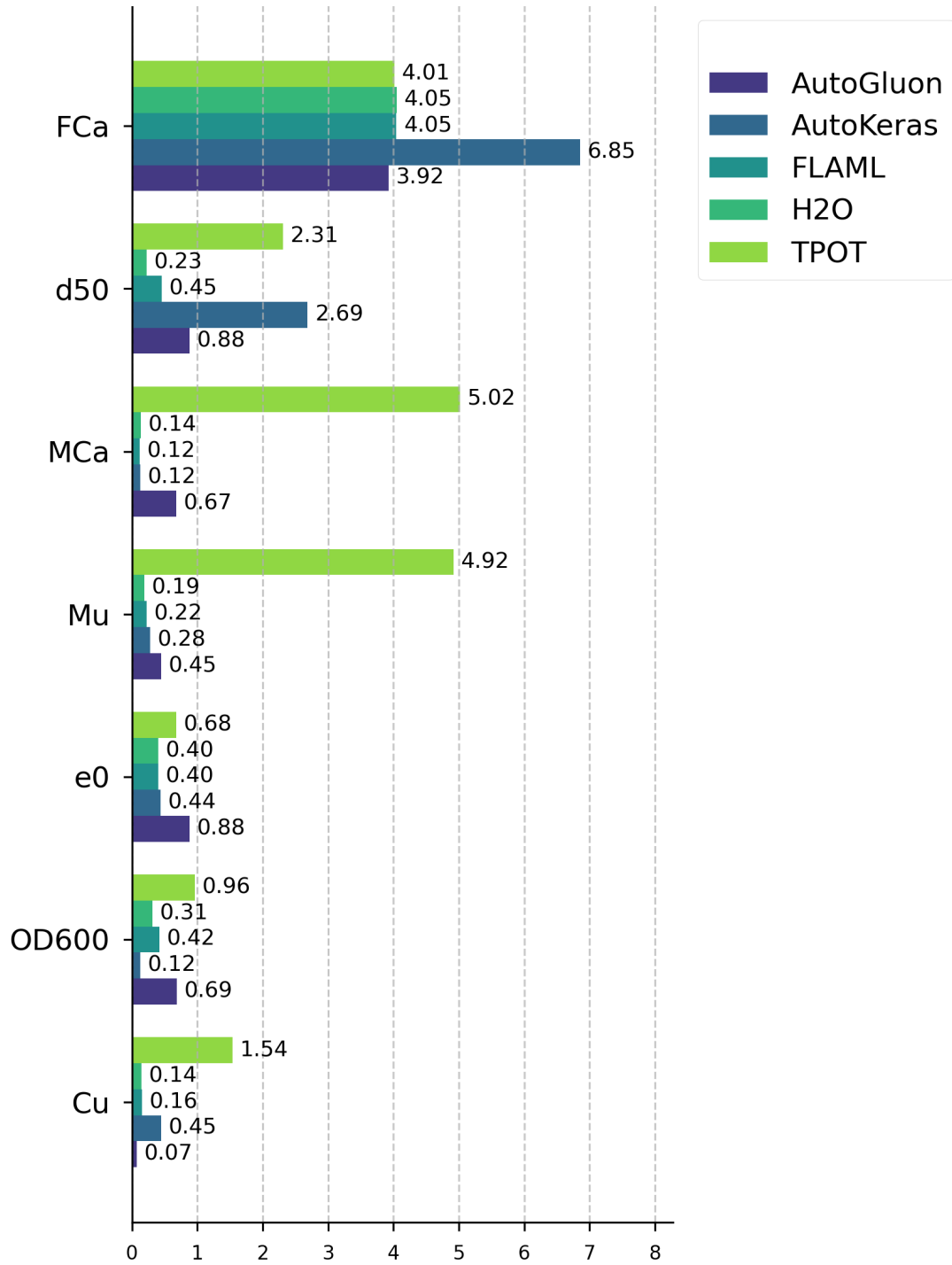
Fonte: Elaborado pelo autor (2025).

Figura 48 – Importância média dos atributos para o conjunto de dados D8.



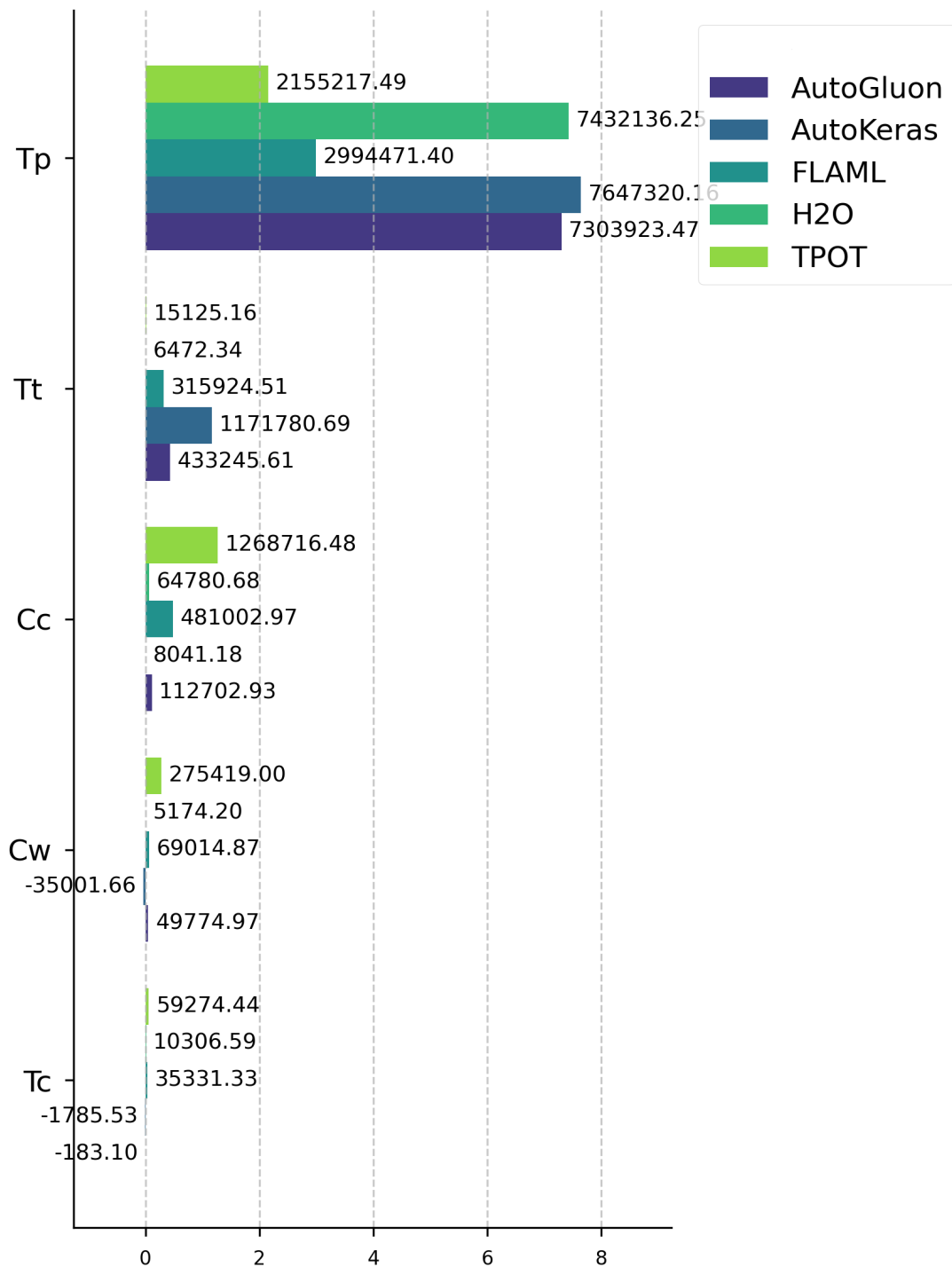
Fonte: Elaborado pelo autor (2025).

Figura 49 – Importância média dos atributos para o conjunto de dados D9.



Fonte: Elaborado pelo autor (2025).

Figura 50 – Importância média dos atributos para o conjunto de dados D10.



Fonte: Elaborado pelo autor (2025).

No conjunto D9 o FLAML obteve o melhor resultado, sendo que o mesmo priorizou e realizou uma seleção balanceada do teor de carbonato de cálcio, do diâmetro médio dos grãos e da turbidez da suspensão bacteriana em relação aos demais atributos, quando comparado com os demais sistemas. Já no conjunto D10, a elevada frequência de seleção da proporção de cimento sugere sua relevância para o bom desempenho do TPOT. Também

vale observar a alta frequência de seleção dos atributos relacionados à resistência de aderência à tração.

As diferenças na hierarquização das variáveis podem ser explicadas pelas relações físicas entre os parâmetros geotécnicos considerados. Umidade, densidade seca, porosidade e plasticidade representam aspectos interdependentes do comportamento do solo. A variação do teor de umidade afeta a estrutura do solo, o arranjo das partículas e o resultado da compactação, refletindo diretamente na densidade. Assim, variáveis distintas podem expressar informações complementares sobre um mesmo processo físico. Esse cenário é típico de conjuntos de dados geotécnicos, nos quais a multicolinearidade é frequente. Alterações induzidas pela umidade afetam simultaneamente densidade, rigidez e sucção, formando grupos de variáveis com significado físico semelhante (129, 130, 131).

Diferentes algoritmos podem selecionar atributos distintos sem comprometer o desempenho preditivo. Conforme destacado por Nguyen et al. (132), bases de dados de materiais frequentemente contêm variáveis correlacionadas, levando os modelos a explorar descritores alternativos para representar o mesmo fenômeno. Além disso, cada algoritmo interpreta de forma distinta as não linearidades e interações entre variáveis. Métodos baseados em árvores tendem a privilegiar atributos que geram divisões locais mais eficientes, enquanto modelos lineares ou regularizados favorecem relações globais mais estáveis. Resultados como os apresentados por Oo et al. (133) mostram que, classificações ordenadas de importância variam significativamente entre algoritmos na presença de preditores correlacionados.

Portanto, as diferenças observadas na seleção de variáveis não indicam necessariamente fragilidade dos sistemas avaliados. Elas refletem a natureza correlacionada dos parâmetros geotécnicos e as diferentes estratégias utilizadas para representar relações não lineares. Mesmo com prioridades distintas, os modelos conseguem capturar os mecanismos fundamentais que controlam a RCS, mantendo desempenho preditivo consistente.

4.3 INTERPRETAÇÕES DOS RESULTADOS

A avaliação conjunta das métricas permite uma interpretação mais consistente do comportamento dos sistemas AutoML. As estatísticas R e R^2 descrevem o grau de ajuste entre valores preditos e observados. Em contraste, RMSE e MAE apresentaram variações reduzidas entre os conjuntos analisados, o que indica estabilidade frente a observações extremas. De forma geral, os sistemas alcançaram resultados satisfatórios para R^2 , R e MAPE. Observou-se ainda que melhorias no ajuste global (R e R^2) estiveram associadas à redução dos erros absolutos (RMSE e MAE) e relativos (MAPE), ou seja, conjuntos com maiores valores de R e R^2 tendem a apresentar menores RMSE, MAE e MAPE (Tabelas 17 a 20).

A métrica MAPE mostrou maior instabilidade quando os valores reais se aproxi-

maram de zero. Esse comportamento foi recorrente nos conjuntos D1, D2, D5 e D6, com maior intensidade em D10. Entre os conjuntos avaliados, D4 apresentou os menores erros relativos médios, embora o AutoKeras tenha mantido desempenho inferior em relação aos demais métodos (Tabelas 17 a 20).

O bom desempenho reportado por Goliatt et al. (58) nos conjuntos D1 a D6, evidencia a importância da escolha adequada de modelos e da calibração sistemática de hiperparâmetros na estimativa da RCS (Tabelas 17 e 18). Em contrapartida, os sistemas AutoML mostraram resultados comparáveis aos modelos originais de estudo dos conjuntos de dados analisados (Tabela 21).

A análise do ID (Tabelas 22 e 23) indica o FLAML como uma alternativa preferencial quando se busca desempenho global consistente em diferentes conjuntos de dados. O TPOT surge como uma opção viável quando se busca equilíbrio entre desempenho e simplicidade. O H2O apresentou resultados expressivos em bases mais heterogêneas, como D1, D3, D4 e D6, com valores elevados de ID, indicando boa adaptação a dados complexos. Considerando o tempo de execução, o AutoGluon mostra-se adequado em aplicações nas quais a rapidez é priorizada em detrimento de ganhos marginais de precisão.

Ao se comparar os resultados resumidos dos IDs (Tabela 23) e a quantidade de diferenças significativas entre os sistemas por conjuntos de dados (Tabela 25), é interessante notar que os conjuntos D4 e D10 apresentaram os maiores valores de ID médio e a maior quantidade de diferenças significativas entre os sistemas.

O teste de Kruskal–Wallis indicou maior quantidade de diferenças significativas entre conjuntos de dados para R , R^2 e MAE, entretanto no teste de Dunn, as métricas RMSE, MAE e MAPE apresentaram quantidades maiores de diferenças significativas entre sistemas AutoML (Figura 34 e Tabela 24). Os resultados indicam que as diferenças entre os sistemas são mais evidentes na magnitude dos erros de predição do que na capacidade de representar a variabilidade dos dados. Essa interpretação é reforçada pela análise estatística não paramétrica, que aponta distinções mais consistentes nos indicadores de erro do que nas métricas associadas ao ajuste global do modelo.

Os resultados do teste de Dunn confirmam que as diferenças observadas não são aleatórias. Em grande parte dos cenários, com significância estatística, os sistemas H2O, FLAML e TPOT formaram um grupo com desempenhos semelhantes com poucas diferenças entre si. Esse comportamento indica desempenho estável desses modelos em diferentes conjuntos de dados (Figuras 36 a 40).

Considere o conjunto D5, analisando a importância das variáveis (Figura 45) e os resultados de desempenho (Tabela 22). Como AutoKeras e FLAML apresentaram o mesmo nível de desempenho, e os atributos priorizados por cada um foram semelhantes, isso sugere que a escolha das variáveis gera forte influência na qualidade preditiva dos sistemas AutoML.

A avaliação de importância das variáveis mostrou um padrão consistente entre os conjuntos analisados. Variáveis diretamente associadas à quantidade de estabilizante (conjuntos D1, D2, D3, D4, D5 e D9) e a compactação dos solos (conjuntos D6 e D7) apresentaram maior influência nas estimativas.

Sob o ponto de vista geotécnico, as associações estatísticas identificadas entre variáveis de entrada e a RCS são compatíveis com mecanismos físicos e químicos amplamente documentados (129, 130, 131, 77, 4, 6, 2, 70, 67, 68, 71, 72, 49, 50, 51, 3, 52, 20, 54). Essa compatibilidade contribui para a confiabilidade dos modelos e reforça sua aplicabilidade prática. A elevada relevância atribuída ao teor de estabilizante, incluindo cimento, cal e GGBS, reflete diretamente os processos de coesão, nos quais reações de hidratação e de caráter pozolânico promovem a formação de produtos ligantes, como o silicato de cálcio hidratado. Esses compostos atuam na conexão entre partículas, resultando em incremento significativo da resistência mecânica.

A cura das amostras também se destacou como variável relevante (conjuntos D3 e D7), o que é fisicamente consistente, pois representa a progressão temporal das reações químicas responsáveis pelo ganho de resistência. À medida que essas reações evoluem, a estrutura ligante torna-se mais coesa e estável. Parâmetros relacionados à compactação (conjuntos D6 e D7) refletem o arranjo das partículas. Estruturas mais densas apresentam menor volume de vazios e maior número de contatos intergranulares, favorecendo a atuação dos agentes ligantes.

Sabe-se que a densidade é um fator importante na predição da RCS de solos estabilizados com cimento. Entretanto, os sistemas que obtiveram os melhores resultados não selecionaram essa variável com frequência (conjunto D3), o que sugere que seria interessante a realização de novos testes para verificar os resultados, caso esse atributo tivesse um maior peso nos modelos de aprendizagem de máquina.

A importância da composição do solo e do índice de plasticidade (conjuntos D1, D2, D5 e D6) está associada à interação entre o solo e o estabilizante. Minerais argilosos oferecem superfícies reativas que contribuem para as reações químicas, mas também exigem maiores teores de água e material estabilizante. Essa relação é intrinsecamente não linear e foi adequadamente capturada pelos modelos. No conjunto D9, estabilizado por MICP, a relevância do teor de carbonato de cálcio está diretamente ligada ao mecanismo predominante de ganho de resistência, no qual a calcita precipitada atua como agente ligante, preenchendo vazios e conectando partículas.

No conjunto D10, todos os sistemas selecionaram com alta frequência os atributos relacionados à resistência de aderência à tração do solo-cimento com as barras GFRP. Esse comportamento é esperado, uma vez que essas variáveis refletem um componente de resistência dos solos, o que está altamente correlacionado com a RCS.

5 CONCLUSÕES

Este trabalho analisou o uso de sistemas de AutoML para a estimativa da RCS de solos estabilizados, a partir de bases experimentais reunidas em diferentes países. Foram avaliados cinco sistemas amplamente difundidos, considerando variações de solo, tipos de estabilização e condições de ensaio. As bases incluíram estabilizantes tradicionais, como cimento e cal, além de materiais alternativos, como cinza de casca de arroz, ligantes geopoliméricos e tratamentos de origem bioquímica. A abrangência geográfica dos dados permitiu examinar a consistência dos modelos em cenários distintos.

Os resultados indicam que o AutoML apresenta desempenho adequado para a predição da RCS, porém, sua aplicação é mais indicada em estudos nos quais a máxima precisão não é o objetivo principal. Abordagens baseadas em modelos específicos, com ajuste manual e otimização direcionada, tendem a produzir resultados superiores, conforme evidenciado por Goliatt et al. (58). Ainda assim, os sistemas de AutoML atendem à proposta de automatizar o processo de modelagem, reduzindo a exigência de conhecimento especializado em aprendizado de máquina. Esses sistemas também oferecem um fluxo padronizado e reproduzível para o tratamento de bases heterogêneas, com menor intervenção nas etapas de seleção de atributos, escolha de algoritmos e ajuste de hiperparâmetros.

Do ponto de vista da engenharia geotécnica, os resultados apresentam implicações relevantes. O uso de AutoML possibilita a geração rápida de modelos preditivos, diminuindo a dependência de campanhas experimentais extensas. Além disso, a aplicação dos sistemas a bases provenientes de diferentes regiões demonstra que padrões coerentes de comportamento solo–estabilizante podem ser identificados de forma automatizada. A comparação direta entre sistemas também fornece subsídios objetivos para a escolha de ferramentas em aplicações futuras de predição da RCS.

Considerando o compromisso entre desempenho e custo computacional, podem ser feitas recomendações específicas. Quando o tempo de execução é um fator limitante, o AutoGluon se mostra adequado para análises exploratórias e estudos preliminares, com resultados obtidos em curto intervalo de tempo. Embora o H2O e o TPOT sejam alternativas robustas em ambientes mais estruturados e o AutoKeras seja voltado à aplicações baseadas em redes neurais, o FLAML se mostrou mais flexível frente à diversidade de bases típicas da engenharia geotécnica sendo a melhor opção quando a prioridade recai sobre desempenho consistente. Essas diretrizes auxiliam na aplicação prática do AutoML em projetos de estabilização de solos.

De forma geral, os resultados indicam que os sistemas AutoML foram capazes de aprender relações coerentes com princípios fundamentais da mecânica dos solos. Assim, as predições obtidas não se restringem a ajustes estatísticos, mas refletem processos físicos relevantes, aumentando o valor interpretativo dos modelos para aplicações em

engenharia. Além disso, vale observar que um único sistema não consegue apresentar bom desempenho para todos os tipos de dados. Um bom exemplo é o AutoKeras, que apesar de obter resultado global inferior, obteve os melhores resultados para os conjuntos D2 e D8 (Tabela 22).

Em linhas gerais, os resultados mostram que os sistemas de AutoML permitem automatizar a predição da RCS, oferecendo maior acessibilidade e reprodutibilidade para estudos geotécnicos. Embora não substituam abordagens específicas e cuidadosamente ajustadas em aplicações que exigem alta precisão, essas ferramentas constituem um recurso útil para acelerar análises e apoiar decisões baseadas em dados no contexto da estabilização de solos.

5.1 LIMITAÇÕES

É necessário considerar limitações inerentes dos sistemas de AutoML, como a dependência da qualidade dos dados, o custo computacional associado a buscas extensas, e a possível perda de nuances físicas que podem ser captadas por engenharia de atributos orientada por especialistas.

O aumento do grau de automação em ciência de dados exige atenção aos seus efeitos colaterais. Entre eles estão a confiança excessiva nos resultados gerados por sistemas automatizados, a ocorrência de falhas discretas e difíceis de identificar, e o reforço de vieses cognitivos associados aos tipos de dados, modelos e interpretações privilegiadas pelas ferramentas disponíveis. No contexto de cooperação entre humanos e inteligência artificial, esses sistemas também modificam a prática profissional. Parte das atividades tradicionais são alteradas e novas funções emergem, na quais as ferramentas deixam de atuar apenas como suporte operacional e passam a assumir um papel mais ativo, funcionando como assistentes no processo analítico (84).

Os ensaios realizados em laboratório apresentam limitações importantes quanto à representação fiel das condições naturais dos solos. O processo de amostragem interfere diretamente nas propriedades medidas. Durante a cravação do amostrador, a retirada do material e a moldagem dos corpos de prova, ocorrem perturbações mecânicas inevitáveis. Evidências experimentais indicam que as amostras não preservam as tensões neutras negativas existentes no estado natural. Essa perda decorre do rearranjo das partículas e da redução das forças associadas à água adsorvida. Soma-se a isso a mudança do estado de tensões, que passa de anisotrópico, típico do campo, para aproximadamente isotrópico nos corpos de prova, alterando o comportamento mecânico observado (67).

O ensaio de compressão simples é aplicável principalmente a solos argilosos, uma vez que não considera o efeito do confinamento. Para outros materiais, existem métodos mais adequados para a avaliação da resistência ao cisalhamento, como os ensaios de cisalhamento direto, palheta e compressão triaxial. O ensaio triaxial é amplamente empregado por

permitir o controle das tensões de confinamento, das condições de drenagem e da taxa de aplicação de carga. Já o ensaio de palheta é especialmente indicado quando se busca estimar parâmetros de resistência diretamente em campo (67, 68).

Com base em resultados experimentais, observa-se que o ensaio de compressão simples apresenta restrições significativas. Seus resultados são sensíveis aos efeitos da amostragem e do armazenamento das amostras. Além disso, o ensaio deve ser conduzido em um curto intervalo de tempo para evitar drenagem. Ensaios executados em altas velocidades tendem a fornecer valores de resistência superiores aos associados às condições reais de carregamento, como aquelas observadas em obras de terra. Enquanto as perturbações reduzem a resistência medida, a velocidade elevada tende a aumentá-la. Esses efeitos atuam em sentidos opostos e podem se compensar parcialmente, o que dificulta a identificação da resistência efetiva do solo por meio desse ensaio (67).

A análise teórica indica que melhorias no desempenho de sistemas de AutoML podem ser obtidas pela ampliação do espaço de busca, pelo aumento do número de iterações dos algoritmos ou pela adoção de estratégias de avaliação mais custosas. Entretanto, tais medidas implicam na elevação do custo computacional. Dessa forma, a relação entre ganho de desempenho e demanda computacional constitui um fator determinante no projeto e na aplicação prática de métodos de AutoML (83).

5.2 TRABALHOS FUTUROS

Pesquisas futuras devem avançar para além da validação interna e investigar a transferência de modelos entre diferentes conjuntos de dados. Neste estudo, a análise foi deliberadamente restrita a modelagens independentes, em que cada base define um problema específico, sem ajuste de um modelo global único. Estratégias de aprendizado por transferência representam um caminho promissor, pois permitem adaptar modelos treinados em bases amplas a novas condições experimentais ou regiões distintas (134, 135). Essa abordagem é especialmente relevante diante da diversidade de solos, estabilizantes e contextos geográficos presentes nos conjuntos D1 a D10.

Trabalhos futuros também podem se beneficiar da inclusão de novos parâmetros físicos e ampliação dos conjuntos de dados disponíveis, principalmente de conjuntos de rochas que neste estudo foi tratado intencionalmente de forma isolada e limitada. Inclusão de outros sistemas de AutoML, conforme sugerido por (82), é promissora para o desenvolvimento de novos estudos.

Como perspectiva para trabalhos futuros, destaca-se a necessidade de desenvolvimento e disponibilização de conjuntos de dados voltados a solos brasileiros ou tropicais, que contemplem informações experimentais associadas a técnicas de aprendizado de máquina. A escassez de bases de dados públicas que atendam simultaneamente a critérios como publicação científica, disponibilização dos dados, aplicação de modelos de regressão e

apresentação de métricas de desempenho limita análises comparativas mais abrangentes. Nesse sentido, recomenda-se a realização de estudos que integrem campanhas experimentais e modelagem preditiva, com a devida padronização e divulgação dos dados, a fim de possibilitar a validação cruzada de modelos e o avanço do conhecimento aplicado a condições geotécnicas representativas do contexto brasileiro.

Estudos futuros podem explorar combinações entre sistemas, visando integrar estratégias complementares e ampliar a eficiência dos modelos AutoML. A análise de tempos de treinamento mais longos pode ainda permitir comparações mais robustas com os resultados apresentados por Goliatt et al. (58).

REFERÊNCIAS

- 1 V Q Tran and L Q Nguyen. Using machine learning technique for designing reinforced lightweight soil. *Journal of Intelligent and Fuzzy Systems*, 43(1):1633–1650, 2022. ISSN 10641246 (ISSN). doi: 10.3233/JIFS-212621.
- 2 Nilo Cesar Consoli, Helena Batista Leon, Mariana da Silva Carretta, João Victor Linch Daronco, and David Eduardo Lourenço. The effects of curing time and temperature on stiffness, strength and durability of sand-environment friendly binder blends. *Soils and Foundations*, 59(5):1428–1439, 2019.
- 3 Woubishet Zewdu Taffese and Kassahun Admassu Abegaz. Artificial intelligence for prediction of physical and mechanical properties of stabilized soil for affordable housing. *Applied Sciences*, 11(16), 2021. ISSN 2076-3417. doi: 10.3390/app11167503. URL <https://www.mdpi.com/2076-3417/11/16/7503>.
- 4 Youchao Zhang, Shuangli Jiang, Dengzhou Quan, Kun Fang, Bo Wang, and Zhiming Ma. Properties of sustainable earth construction materials: A state-of-the-art review. *Sustainability*, 16(2):670, 2024.
- 5 E.U. Eyo and S.J. Abbey. Machine learning regression and classification algorithms utilised for strength prediction of opc/by-product materials improved soils. *Construction and Building Materials*, 284:122817, 2021. ISSN 0950-0618. doi: 10.1016/j.conbuildmat.2021.122817. URL <https://www.sciencedirect.com/science/article/pii/S0950061821005778>.
- 6 Leonardo Behak. Soil stabilization with rice husk ash. In Amanullah and Shah Fahad, editors, *Rice*, chapter 3. IntechOpen, Rijeka, 2017. doi: 10.5772/66311. URL <https://doi.org/10.5772/66311>.
- 7 Sepehr Soleimani, Shabnam Rajaei, Pengcheng Jiao, Arash Sabz, and Sina Soheilinia. New prediction models for unconfined compressive strength of geopolymer stabilized soil using multi-gen genetic programming. *Measurement*, 113:99–107, 2018.
- 8 A Gajurel, P S Mukherjee, and B Chittoori. Estimating Optimal Additive Content for Soil Stabilization Using Machine Learning Methods. In Meehan M.L., Kumar S., Pando M.A., and Coe J.T., editors, *8th International Conference on Case Histories in Geotechnical Engineering: Geotechnical Materials, Modeling, and Testing, Geo-Congress 2019*, number GSP 310, pages 662–672, Dept. of Civil Engineering, Boise State Univ., 1910 W. University Dr., Boise, ID 83725, United States, 2019. American Society of Civil Engineers Reston, VA, American Society of Civil Engineers (ASCE). ISBN 08950563 (ISSN); 9780784482124 (ISBN); 9780784482124 (ISBN). doi: 10.1061/9780784482124.067.
- 9 Genbao Zhang, Changfu Chen, Junbo Sun, Kefei Li, Fan Xiao, Yufei Wang, Mengcheng Chen, Jizhuo Huang, and Xiangyu Wang. Mixture optimisation for cement-soil mixtures with embedded gfrp tendons. *Journal of Materials Research and Technology*, 18:611–628, 2022.
- 10 S. Mohanty, N. Roy, S. P. Singh, and P. Sihag. Estimating the strength of stabilized dispersive soil with cement clinker and fly ash. *Geotech Geol Eng*, 37:2915–2926, 2019.

- 11 G. Zhang, Z. Ding, Y. Wang, G. Fu, Y. Wang, C. Xie, Y. Zhang, X. Zhao, X. Lu, and X. Wang. Performance prediction of cement stabilized soil incorporating solid waste and propylene fiber. *Materials*, 15:4250, 2022.
- 12 H. Chen and L. Zhang. A machine learning-based method for predicting end-bearing capacity of rock-socketed shafts. *Rock Mech Rock Eng*, 55:1743–1757, 2022.
- 13 C. Kayadelen, G. Altay, and Y. Önal. Numerical simulation and novel methodology on resilient modulus for traffic loading on road embankment. *Int J Pavement Eng*, 2021.
- 14 H.-B. Ly and B. T. Phaim. Soil unconfined compressive strength prediction using random forest (rf) machine learning model. *Open Constr Build Technol J*, 14: 278–285, 2020.
- 15 L. Van Giap, T. A. Pham, and H.-L. Vu. Application of random forest model to determine unconfined compressive strength of the soil-cement mixture, 2022. http://dx.doi.org/10.1007/978-981-16-7160-9_179.
- 16 S. Suman, M. Mahamaya, and S. K. Das. Prediction of maximum dry density and unconfined compressive strength of cement stabilised soil using artificial intelligence techniques. *Int J Geosynth Ground Eng*, 2, 2016.
- 17 J. Tinoco, A. Alberto, P. J. V. Oliveira, L. Lemos, and A. G. Correia. Data-driven approach to predict unconfined compressive strength of laboratory soil stabilized with cementitious binders. In *17th European Conference on Soil Mechanics and Geotechnical Engineering, ECSMGE 2019*, 2019. URL <http://dx.doi.org/10.32075/17ECSMGE-2019-0598>.
- 18 J. Tinoco, A. Alberto, P. da Venda, A. Gomes Correia, and L. Lemos. A novel approach based on soft computing techniques for unconfined compressive strength prediction of soil cement mixtures. *Neural Comput Appl*, 32:8985–8991, 2020.
- 19 F. E. Jalal, Y. Xu, M. Iqbal, M. F. Javed, and B. Jamhiri. Predictive modeling of swell strength of expansive soils using artificial intelligence approaches: Ann, anfis and gep. *J Environ Manag*, 289, 2021.
- 20 H.-L. Wang and Z.-Y. Yin. Unconfined compressive strength of bio-cemented sand: state-of-the-art review and MEP-MC-based model development. *Journal of Cleaner Production*, 315, 2021. ISSN 09596526 (ISSN). doi: 10.1016/j.jclepro.2021.128205.
- 21 V.-N. Pham, E. Oh, and D. E. Ong. Effects of binder types and other significant variables on the unconfined compressive strength of chemical-stabilized clayey soil using gene-expression programming. *Neural Comput Appl*, 34:9103–9121, 2022.
- 22 V. Q. Tran. Selection of single machine learning model for designing compressive strength of stabilized soil containing lime, cement and bitumen. *J Intell Fuzzy Systems*, pages 1–18, 2023.
- 23 P. Guo, W. Meng, M. Xu, V. C. Li, and Y. Bao. Predicting mechanical properties of high-performance fiber-reinforced cementitious composites by integrating micromechanics and machine learning. *Materials*, 14, 2021.

- 24 H. Tao, M. Habib, I. Aljarah, H. Faris, H. A. Afan, and Z. M. Yaseen. An intelligent evolutionary extreme gradient boosting algorithm development for modeling scour depths under submerged weir. *Inform Sci*, 570:172–184, 2021.
- 25 X. Ding, M. Amiri, and M. Hasanipanah. Enhancing shear strength predictions of rocks using a hierarchical ensemble model. *Sci Rep*, 14:20268, 2024.
- 26 A. Malik, M. K. Saggi, S. Rehman, H. Sajjad, S. Inyurt, A. S. Bhatia, A. A. Farooque, A. Y. Oudah, and Z. M. Yaseen. Deep learning versus gradient boosting machine for pan evaporation prediction. *Eng Appl Comput Fluid Mech*, 16:570–587, 2022.
- 27 I. Faridmehr, M. Shariq, V. Plevris, and N. Aalimahmoody. Novel hybrid informational model for predicting the creep and shrinkage deflection of reinforced concrete beams containing ggbfs. *Neural Comput Appl*, pages 1–17, 2022.
- 28 H. Tao, S. M. Awadh, S. Q. Salih, S. S. Shafik, and Z. M. Yaseen. Integration of extreme gradient boosting feature selection approach with machine learning models: application of weather relative humidity prediction. *Neural Comput Appl*, 34:515–533, 2022.
- 29 A. R. Ghanizadeh, N. Heidarabadizadeh, M. Bayat, and V. Khalifeh. Modeling of unconfined compressive strength and young’s modulus of lime and cement stabilized clayey subgrade soil using evolutionary polynomial regression (epr). *Int J Min Geo-Eng*, 2022.
- 30 S. Shamshirband, A. Tavakkoli, C. B. Roy, S. Motamedi, K.-I. Song, R. Hashim, and S. M. Islam. Hybrid intelligent model for approximating unconfined compressive strength of cement-based bricks with odd-valued array of peat content (0%-29%). *Powder Technol*, 284:560–570, 2015.
- 31 G. Zhang, C. Chen, K. Li, F. Xiao, J. Sun, Y. Wang, and X. Wang. Multi-objective optimisation design for gfrp tendon reinforced cemented soil. *Constr Build Mater*, 320:126297, 2022.
- 32 H.-A. Le, T.-A. Nguyen, D.-D. Nguyen, and I. Prakash. Prediction of soil unconfined compressive strength using artificial neural network model. *Vietnam J Earth Sci*, 42: 255–264, 2020.
- 33 A. Q. Ngo, L. Q. Nguyen, and V. Q. Tran. Developing interpretable machine learning shapley additive explanations model for unconfined compressive strength of cohesive soils stabilized with geopolymer. *Plos One*, 18:e0286950, 2023.
- 34 H. Güllü and H. İ. Fedakar. On the prediction of unconfined compressive strength of silty soil stabilized with bottom ash, jute and steel fibers via artificial intelligence. *Geomech Eng*, 12:441–464, 2017.
- 35 J. Tinoco, A. A. S. Correia, and P. da Venda. Soil–cement mixtures reinforced with fibers: A data-driven approach for mechanical properties prediction. *Appl Sci (Switzerland)*, 11, 2021.
- 36 N. Tiwari and N. Satyam. Coupling effect of pond ash and polypropylene fiber on strength and durability of expansive soil subgrades: An integrated experimental and machine learning approach. *J Rock Mech Geotech Eng*, 13:1101–1112, 2021.

- 37 K C Onyelowe, M Iqbal, F E Jalal, M E Onyia, and I C Onuoha. Application of 3-algorithm ANN programming to predict the strength performance of hydrated-lime activated rice husk ash treated soil. *Multiscale and Multidisciplinary Modeling, Experiments and Design*, 4(4):259–274, 2021. ISSN 25208179 (ISSN). doi: 10.1007/s41939-021-00093-7.
- 38 X. Yang, J. Guan, L. Ding, Z. You, V. C. Lee, M. R. M. Hasan, and X. Cheng. Research and applications of artificial neural network in pavement engineering: a state-of-the-art review. *J Traffic Transp Eng (Engl Edition)*, 8:1000–1021, 2021.
- 39 H.-G. Hoang and T.-A. Nguyen. An artificial intelligence approach to predict the resilient modulus of subgrade pavement or unbound material, 2022. http://dx.doi.org/10.1007/978-981-16-7160-9_177.
- 40 A. Sezer, G. İ. Sezer, A. Mardani-Aghabaglou, and S. Altun. Prediction of mechanical and penetrability properties of cement-stabilized clay exposed to sulfate attack by use of soft computing methods. *Neural Comput Appl*, 32:16707–16722, 2020.
- 41 K. Oda. Method for objectively determining design strength with artificial neural network. In *16th Asian Regional Conference on Soil Mechanics and Geotechnical Engineering, ARC 2019*, Osaka Sangyo University, Japan, 2020.
- 42 H. Yao, Z. Xu, Y. Hou, Q. Dong, P. Liu, Z. Ye, X. Pei, M. Oeser, L. Wang, and D. Wang. Advanced industrial informatics towards smart, safe and sustainable roads: A state of the art. *J Traffic Transp Eng (Engl Edition)*, 2023.
- 43 Asskar Janalizadeh Choobbasti and Saman Soleimani Kutanaei. Microstructure characteristics of cement-stabilized sandy soil using nanosilica. *Journal of Rock Mechanics and Geotechnical Engineering*, 9(5):981–988, 2017. ISSN 1674-7755. doi: <https://doi.org/10.1016/j.jrmge.2017.03.015>.
- 44 Xin Kang, Louis Ge, Gi-Chun Kang, and Christopher Mathews. Laboratory investigation of the strength, stiffness, and thermal conductivity of fly ash and lime kiln dust stabilised clay subgrade materials. *Road Materials and Pavement Design*, 16(4):928–945, 2015. doi: 10.1080/14680629.2015.1028970.
- 45 National Lime Association. Technical brief: Mixture design and testing procedures for lime stabilized soil. Technical report, National Lime Association, 2006.
- 46 Portland Cement Association. Soil cement laboratory handbook. Technical report, Portland Cement Association, 1992.
- 47 U.S. Army. Soil stabilization for pavements. Technical Report TM 5-882-14/AFM 32-1019, U.S. Army, 1994.
- 48 A Gajurel, B Chittoori, P S Mukherjee, and M Sadegh. Machine learning methods to map stabilizer effectiveness based on common soil properties. *Transportation Geotechnics*, 27:100506, 2021. ISSN 22143912 (ISSN). doi: 10.1016/j.trgeo.2020.100506. URL <https://www.sciencedirect.com/science/article/pii/S2214391220303949>.

- 49 Huong Thi Thanh Ngo, Tuan Anh Pham, Huong Lan Thi Vu, and Loi Van Giap. Application of artificial intelligence to determined unconfined compressive strength of cement-stabilized soil in vietnam. *Applied Sciences*, 11(4), 2021. ISSN 2076-3417. doi: 10.3390/app11041949. URL <https://www.mdpi.com/2076-3417/11/4/1949>.
- 50 Akash Priyadarshiee, Sunayana Chandra, Deepak Gupta, and Vikas Kumar. Neural models for unconfined compressive strength of kaolin clay mixed with pond ash, rice husk ash and cement. *Journal of Soft Computing in Civil Engineering*, 4(2):85–102, 2020.
- 51 Ruhul Amin Mozumder and Aminul Islam Laskar. Prediction of unconfined compressive strength of geopolymer stabilized clayey soil using artificial neural network. *Computers and Geotechnics*, 69:291–300, 2015. ISSN 0266-352X. doi: 10.1016/j.compgeo.2015.05.021. URL <https://www.sciencedirect.com/science/article/pii/S0266352X15001287>.
- 52 A Tabarsa, N Latifi, A Osouli, and Y Bagheri. Unconfined compressive strength prediction of soils stabilized using artificial neural networks and support vector machines. *Frontiers of Structural and Civil Engineering*, 15(2):520–536, Apr 2021. ISSN 20952430 (ISSN). doi: 10.1007/s11709-021-0689-9.
- 53 Arsalan Mahmoodzadeh, Mokhtar Mohammadi, Hawkar Hashim Ibrahim, Sazan Nariman Abdulhamid, Sirwan Ghafoor Salim, Hunar Farid Hama Ali, and Mohammed Kamal Majeed. Artificial intelligence forecasting models of uniaxial compressive strength. *Transportation Geotechnics*, 27:100499, 2021. ISSN 2214-3912. doi: <https://doi.org/10.1016/j.trgeo.2020.100499>.
- 54 Genbao Zhang, Changfu Chen, Junbo Sun, Kefei Li, Fan Xiao, Yufei Wang, Mengcheng Chen, Jizhuo Huang, and Xiangyu Wang. Mixture optimisation for cement-soil mixtures with embedded gfrp tendons. *Journal of Materials Research and Technology*, 18:611–628, 2022.
- 55 Hassan Eldeeb, Mohamed Maher, Radwa Elshawi, and Sherif Sakr. Automlbench: A comprehensive experimental evaluation of automated machine learning frameworks. *Expert Systems with Applications*, 243:122877, 2024. ISSN 0957-4174. doi: <https://doi.org/10.1016/j.eswa.2023.122877>. URL <https://www.sciencedirect.com/science/article/pii/S0957417423033791>.
- 56 Xin He, Kaiyong Zhao, and Xiaowen Chu. Automl: A survey of the state-of-the-art. *Knowledge-based systems*, 212:106622, 2021.
- 57 Frank Hutter, Lars Kotthoff, and Joaquin Vanschoren. *Automated Machine Learning: methods, systems, challenges*. Springer Cham, 2019. ISBN 978-3-030-05318-5. doi: <https://doi.org/10.1007/978-3-030-05318-5>.
- 58 Leonardo Goliatt, Haydar Abdulameer Marhoon, Zaher Mundher Yaseen, Salim Heddami, Ahmed W. Al Zand, Bijay Halder, Mou Leong Tan, Zulfaqar Sa’adi, Iman Ahmadianfar, and Salah Elsayed. An evolutionary optimized automated machine learning approach to soil unconfined compressive strength prediction for sustainable transportation infrastructure. *Transportation Geotechnics*, 52:101550, 2025. ISSN 2214-3912. doi: <https://doi.org/10.1016/j.trgeo.2025.101550>. URL <https://www.sciencedirect.com/science/article/pii/S2214391225000698>.

- 59 Lei Wang. Automated machine learning techniques for estimating the unconfined compressive strength of soil stabilization. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 19:6347–6359, 01 2025. doi: 10.1007/s12008-024-02209-y.
- 60 Deivid Campos, Bruno da Silva Macêdo, Zainab Al-Khafaji, Melike Aktaş Bozkurt, İhsan Erdem Kayral, Tiago Silveira Gontijo, Matteo Bodini, Camila M Saporetti, and Leonardo Goliatt. Assessment of automl frameworks for predicting compressive and flexural strength of recycled aggregate concrete. *Materials Today Sustainability*, 31:101200, 2025. doi: 10.1016/j.mtsust.2025.101200.
- 61 N Kardani, A Zhou, S.-L. Shen, and M Nazem. Estimating unconfined compressive strength of unsaturated cemented soils using alternative evolutionary approaches. *Transportation Geotechnics*, 29:100591, 2021. ISSN 22143912 (ISSN). doi: 10.1016/j.trgeo.2021.100591.
- 62 T. Q. Ngo, L. Q. Nguyen, and V. Q. Tran. Novel hybrid machine learning models including support vector machine with meta-heuristic algorithms in predicting unconfined compressive strength of organic soils stabilised with cement and lime. *International Journal of Pavement Engineering*, 24(16):3191–3205, 2023. doi: 10.1080/10298436.2022.2136374.
- 63 K. C. Onyelowe, A. A. B. Moghal, A. Ebid, A. U. Rehman, S. Hanandeh, and V. Priyan. Estimating the strength of soil stabilized with cement and lime at optimal compaction using ensemble-based multiple machine learning. *Scientific Reports*, 14(1):17400, 2024. doi: 10.1038/s41598-024-66295-4.
- 64 Q. Yao, X. Chen, J. Zhang, and F. Li. Hybrid xgboost models for predicting unconfined compressive strength of stabilized soils. *Construction and Building Materials*, 414:133220, 2024. doi: 10.1016/j.conbuildmat.2024.133220.
- 65 I.-B. Teodoru, M. M. Voinea, A. D. Ene, G. O. Mihai, A. M. Dumitrescu, D. G. Radu, A. P. Chivu, and I. M. Radulescu. Prediction of unconfined compressive strength in cement-treated soils: A machine learning approach. *Applied Sciences*, 15(13):7022, 2025. doi: 10.3390/app15137022.
- 66 T. R. Kleinert. *Estabilização de solos tropicais com cal e impactos no dimensionamento mecânico-empírico de pavimentos*. Tese de doutorado, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brasil, 2021. Programa de Pós-Graduação em Engenharia Civil: Construção e Infraestrutura.
- 67 Carlos de Sousa Pinto. *Curso Básico de Mecânica dos Solos em 16 Aulas*. Oficina de Textos, São Paulo, 3 edition, 2006. ISBN 978-85-86238-51-2.
- 68 Homero Pinto Caputo, Armando Negreiros Caputo, and J. Martinho de A. Rodrigues. *Mecânica dos solos e suas aplicações, volume 1: fundamentos*. LTC, Rio de Janeiro, 7 edition, 2015. ISBN 978-85-216-3004-3.
- 69 Katia Vanessa Bicalho. Notas de aula de mecânica dos solos i, 2021.
- 70 Braja M. Das and Khaled Sobhan. *Fundamentos de engenharia geotécnica*. Cengage Learning, São Paulo, 8 edition, 2014. ISBN 978-85-221-1824-3.

- 71 Robert D. Holtz, William D. Kovacs, and Thomas C. Sheahan. *An Introduction to Geotechnical Engineering*. Pearson, Upper Saddle River, 2 edition, 2011.
- 72 James K. Mitchell and Kenichi Soga. *Fundamentals of Soil Behavior*. Wiley, Hoboken, 3 edition, 2005.
- 73 Departamento Nacional de Estradas de Rodagem (DNER). Dner-me 202/94: Solo-cimento — moldagem e cura de corpos-de-prova cilíndricos. Norma técnica, DNER, Rio de Janeiro, Brasil, 1994.
- 74 Departamento Nacional de Estradas de Rodagem (DNER). Dner-me 201/94: Solo-cimento — compressão axial de corpos-de-prova cilíndricos. Norma técnica, DNER, Rio de Janeiro, Brasil, 1994.
- 75 Sravan Muguda, Paul Hughes, Charles Augarde, Céline Perlot, Agostino Walter Bruno, and Domenico Gallipoli. Cross-linking of biopolymers for stabilizing earthen construction materials. *Building Research and Information*, 50:502–514, 01 2022. doi: 10.1080/09613218.2021.2001304.
- 76 Liangbo Ying, Mengqi Xu, Jiale Luo, and Wei Wang. Laboratory investigation and mechanical evaluation on xanthan gum-reinforced clay: Unconfined compression test, triaxial shear test, and microstructure characterization. *Buildings*, 15(14), 2025. ISSN 2075-5309. doi: 10.3390/buildings15142438. URL <https://www.mdpi.com/2075-5309/15/14/2438>.
- 77 J. K. Mitchell. Soil improvement – state-of-the-art report. In *Proceedings of the 10th International Conference on Soil Mechanics and Foundation Engineering*, volume 4, pages 509–565, Stockholm, 1981. International Society for Soil Mechanics and Geotechnical Engineering.
- 78 A. Muthadhi and S. Kothandaraman. Optimum production conditions for reactive rice husk ash. *Materials and Structures*, 43(9):1303–1315, 2010. ISSN 1871-6873. doi: 10.1617/s11527-010-9581-0. URL <https://doi.org/10.1617/s11527-010-9581-0>.
- 79 M. Terashi and I. Juran. Ground improvement – state of the art. In *ISRM International Symposium*, 2000.
- 80 Joel Grus. *Data Science from Scratch: First Principles with Python*. O’Reilly, Beijing, 2015. ISBN 978-1-4919-0142-7. URL <https://learning.oreilly.com/library/view/data-science-from/9781492041122/>.
- 81 Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. URL <http://www.deeplearningbook.org>. Book in preparation for MIT Press.
- 82 Mitra Baratchi, Can Wang, Steffen Limmer, Jan van Rijn, Holger Hoos, Thomas Bäck, and Markus Olhofer. Automated machine learning: past, present and future. *Artificial Intelligence Review*, 57, 04 2024. doi: 10.1007/s10462-024-10726-1.
- 83 Zhenqian Shen, Yongqi Zhang, Lanning Wei, Huan Zhao, and Quanming Yao. Automated machine learning: From principles to practices, 2024. URL <https://arxiv.org/abs/1810.13306>.

- 84 Tijl De Bie, Luc De Raedt, José Hernández-Orallo, Holger H. Hoos, Padhraic Smyth, and Christopher K. I. Williams. Automating data science. *Commun. ACM*, 65(3): 76–87, February 2022. ISSN 0001-0782. doi: 10.1145/3495256. URL <https://doi.org/10.1145/3495256>.
- 85 Jesmeen Hoque, J. Hossen, Md Shohel Sayeed, C.K. Ho, K. Tawsif, Md. Armanur Rahman, and Md Hossain. A survey on cleaning dirty data using machine learning paradigm for big data analytics. *Indonesian Journal of Electrical Engineering and Computer Science*, 10:1234–1243, 06 2018. doi: 10.11591/ijeecs.v10.i3.pp1234-1243.
- 86 Samuel Stocksieker, Denys Pommeret, and Arthur Charpentier. Data augmentation for imbalanced regression. In Francisco Ruiz, Jennifer Dy, and Jan-Willem van de Meent, editors, *Proceedings of The 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pages 7774–7799. PMLR, 25–27 Apr 2023. URL <https://proceedings.mlr.press/v206/stocksieker23a.html>.
- 87 Bartosz Krawczyk. Learning from imbalanced data: Open challenges and future directions. *Progress in Artificial Intelligence*, 5, 04 2016. doi: 10.1007/s13748-016-0094-0.
- 88 Tom Ching LingChen, Ava Khonsari, Amirreza Lashkari, Mina Rafi Nazari, Jaspreet Singh Sambee, and Mario A. Nascimento. Uniformaugmt: A search-free probabilistic data augmentation approach, 2020. URL <https://arxiv.org/abs/2003.14348>.
- 89 M. Dash and H. Liu. Feature selection for classification. *Intelligent Data Analysis*, 1(1):131–156, 1997. ISSN 1088-467X. doi: [https://doi.org/10.1016/S1088-467X\(97\)00008-5](https://doi.org/10.1016/S1088-467X(97)00008-5). URL <https://www.sciencedirect.com/science/article/pii/S1088467X97000085>.
- 90 Hugo Jair Escalante, Manuel Montes, and Luis Enrique Sucar. Particle swarm model selection. *J. Mach. Learn. Res.*, 10:405–440, June 2009. ISSN 1532-4435.
- 91 Ron Kohavi and George H. John. Automatic parameter selection by minimizing estimated error. In *Proceedings of the Twelfth International Conference on International Conference on Machine Learning, ICML'95*, page 304–312, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc. ISBN 1558603778.
- 92 Arber Zela, Aaron Klein, Stefan Falkner, and Frank Hutter. Towards automated deep learning: Efficient joint neural architecture and hyperparameter search, 2018. URL <https://arxiv.org/abs/1807.06906>.
- 93 Hanwen Liang, Shifeng Zhang, Jiacheng Sun, Xingqiu He, Weiran Huang, Kechen Zhuang, and Zhenguo Li. Darts+: Improved differentiable architecture search with early stopping, 2020. URL <https://arxiv.org/abs/1909.06035>.
- 94 Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. Hyperband: A novel bandit-based approach to hyperparameter optimization, 2018. URL <https://arxiv.org/abs/1603.06560>.

- 95 Barret Zoph and Quoc Le. Neural architecture search with reinforcement learning. In *International Conference on Learning Representations*, 2017. URL <https://openreview.net/forum?id=r1Ue8Hcxg>.
- 96 Yi-Qi Hu, Yang Yu, and Zhi-Hua Zhou. Experienced optimization with reusable directional model for hyper-parameter search. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence, IJCAI'18*, page 2276–2282. AAAI Press, 2018. ISBN 9780999241127.
- 97 Peter Bajcsy, Michael Majurski, Thomas E. Cleveland IV, Manuel Carrasco, and Walid Keyrouz. Characterization of ai model configurations for model reuse. In Leonid Karlinsky, Tomer Michaeli, and Ko Nishino, editors, *Computer Vision – ECCV 2022 Workshops*, pages 454–469, Cham, 2023. Springer Nature Switzerland. ISBN 978-3-031-25069-9.
- 98 Nick Erickson, Jonas Mueller, Alexander Shirkov, Hang Zhang, Pedro Larroy, Mu Li, and Alexander Smola. Autogluon-tabular: Robust and accurate automl for structured data. *arXiv preprint arXiv:2003.06505*, 2020.
- 99 Haifeng Jin, Qingquan Song, and Xia Hu. Auto-keras: An efficient neural architecture search system. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 1946–1956, 2019.
- 100 Chi Wang, Qingyun Wu, Markus Weimer, and Erkang Zhu. Flaml: A fast and lightweight automl library. In *MLSys*, 2021.
- 101 Erin LeDell, Sebastien Poirier, et al. H2o automl: Scalable automatic machine learning. In *Proceedings of the AutoML Workshop at ICML*, volume 2020, page 24, 2020.
- 102 Randal S Olson and Jason H Moore. Tpot: A tree-based pipeline optimization tool for automating machine learning. In *Workshop on automatic machine learning*, pages 66–74. PMLR, 2016.
- 103 ASTM International. Standard practice for classification of soils for engineering purposes (unified soil classification system), 2006. ASTM Volume 04.08: Soil And Rock (I): D420 – D5876/D5876m.
- 104 Mustafa Aytakin and Evin Nas. Soil stabilization with lime and cement. *Teknik Dergi/Technical Journal of Turkish Chamber of Civil Engineers*, 9:1573–1585, 12 1998.
- 105 F. G. Bell. Lime stabilisation of clay soils. *Bulletin of the International Association of Engineering Geology – Bulletin de l'Association Internationale de Géologie de l'Ingénieur*, 39(1):67–74, 1989. ISSN 1435-9537. doi: 10.1007/BF02592537. URL <https://doi.org/10.1007/BF02592537>.
- 106 Van Stephan Burroughs. Quantitative criteria for the selection and stabilisation of soils for rammed earth wall construction. Master's thesis, University of New South Wales, 2001.

- 107 Khelifa Harichane, Mohamed Ghrici, Said Kenai, and Khaled Grine. Use of natural pozzolana and lime for stabilization of cohesive soils. *Geotechnical and Geological Engineering*, 29(5):759–769, 2011. ISSN 1573-1529. doi: 10.1007/s10706-011-9415-z. URL <https://doi.org/10.1007/s10706-011-9415-z>.
- 108 Khelifa Harichane, Mohamed Ghrici, Khebizi Wiam, and Hanifi Missoum. Effect of the combination of lime and natural pozzolana on the durability of clayey soils. *Electronic Journal of Geotechnical Engineering*, 15:1194–1209, 01 2010.
- 109 Hayder Hasan, Liet Dang, Hadi Khabbaz, Behzad Fatahi, and Sergei Terzaghi. Remediation of expansive soils using agricultural waste bagasse ash. *Procedia Engineering*, 143:1368–1375, 2016. ISSN 1877-7058. doi: <https://doi.org/10.1016/j.proeng.2016.06.161>. URL <https://www.sciencedirect.com/science/article/pii/S1877705816306087>. Advances in Transportation Geotechnics III.
- 110 B.R. Phanikumar and E. Ramanjaneya Raju. Compaction and strength characteristics of an expansive clay stabilised with lime sludge and cement. *Soils and Foundations*, 60(1):129–138, 2020. ISSN 0038-0806. doi: <https://doi.org/10.1016/j.sandf.2020.01.007>. URL <https://www.sciencedirect.com/science/article/pii/S0038080620300160>.
- 111 Abu Siddique and M. Alomgir Hossain. Effects of lime stabilisation on engineering properties of an expansive soil for use in road construction. *Journal of Society for Transportation and Traffic Studies*, 2(4):1–8, 2011.
- 112 Pranshoo Solanki, Najj N. Khoury, and Musharraf M. Zaman. Engineering properties of stabilized subgrade soils for implementation of the aashto 2002 pavement design guide. Technical report, Norman, Oklahoma, 2009.
- 113 Asres Getnet and Shamshad Alam. Effect of lime and fly ash on load bearing capacity of expansive clay soil. *International Journal of Innovative Technology and Exploring Engineering*, 8(11):554–563, September 2019. doi: 10.35940/ijitee.k1573.0881119. URL <http://dx.doi.org/10.35940/ijitee.K1573.0881119>.
- 114 Leon A. van Paassen, Ranajit Ghose, Thomas J. M. van der Linden, Wouter R. L. van der Star, and Mark C. M. van Loosdrecht. Quantifying biomediated ground improvement by ureolysis: Large-scale biogROUT experiment. *Journal of Geotechnical and Geoenvironmental Engineering*, 136(12):1721–1728, 2010. doi: 10.1061/(ASCE)GT.1943-5606.0000382.
- 115 Liang Cheng, Ralf Cord-Ruwisch, and Mohamed A. Shahin. Cementation of sand soil by microbially induced calcite precipitation at various degrees of saturation. *Canadian Geotechnical Journal*, 50(1):81–90, 2013. doi: 10.1139/cgj-2012-0023.
- 116 L. CHENG, M.A. SHAHIN, and R. CORD-RUWISCH. Bio-cementation of sandy soil using microbially induced carbonate precipitation for marine environments. *Géotechnique*, 64(12):1010–1013, 12 2014. ISSN 0016-8505. doi: 10.1680/geot.14.T.025.
- 117 Liang Cheng, Mohamed A. Shahin, and Donovan Mujah. Influence of key environmental conditions on microbially induced cementation for soil stabilization.

- Journal of Geotechnical and Geoenvironmental Engineering*, 143(1):04016083, 2017. doi: 10.1061/(ASCE)GT.1943-5606.0001586.
- 118 Qian Zhao, Lin Li, Chi Li, Mingdong Li, Farshad Amini, and Huanzhen Zhang. Factors affecting improvement of engineering properties of micp-treated soil catalyzed by bacteria and urease. *Journal of Materials in Civil Engineering*, 26(12):04014094, 2014. doi: 10.1061/(ASCE)MT.1943-5533.0001013.
- 119 A. Mahawish, A. Bouazza, and W. P. Gates. Effect of particle size distribution on the bio-cementation of coarse aggregates. *Acta Geotechnica*, 13:1019–1025, 2018. ISSN 1861-1125. doi: 10.1007/s11440-017-0604-7.
- 120 Aamir Mahawish, Abdelmalek Bouazza, and Will P. Gates. Unconfined compressive strength and visualization of the microstructure of coarse sand subjected to different biocementation levels. *Journal of Geotechnical and Geoenvironmental Engineering*, 145(8):04019033, 2019. doi: 10.1061/(ASCE)GT.1943-5606.0002066.
- 121 Donovan Mujah, Liang Cheng, and Mohamed A. Shahin. Microstructural and geomechanical study on biocemented sand for optimization of micp process. *Journal of Materials in Civil Engineering*, 31(4):04019025, 2019. doi: 10.1061/(ASCE)MT.1943-5533.0002660.
- 122 K. Wen, Y. Li, S. Liu, et al. Development of an improved immersing method to enhance microbial induced calcite precipitation treated sandy soil through multiple treatments in low cementation media concentration. *Geotechnical and Geological Engineering*, 37:1015–1027, 2019. ISSN 0960-3182. doi: 10.1007/s10706-018-0669-6. URL <https://doi.org/10.1007/s10706-018-0669-6>.
- 123 Yang Xiao, Xiang He, T. Matthew Evans, Armin W. Stuedlein, and Hanlong Liu. Unconfined compressive and splitting tensile strength of basalt fiber-reinforced biocemented sand. *Journal of Geotechnical and Geoenvironmental Engineering*, 145(9):04019048, 2019. doi: 10.1061/(ASCE)GT.1943-5606.0002108.
- 124 Ashkan Nafisi, Douglas Mocelin, Brina M. Montoya, and Shane Underwood. Tensile strength of sands treated with microbially induced carbonate precipitation. *Canadian Geotechnical Journal*, 57(10):1611–1616, 2020. doi: 10.1139/cgj-2019-0230.
- 125 Yuze Wang, Kenichi Soga, Jason T. Dejong, and Alexandre J. Kabla. A microfluidic chip and its use in characterising the particle-scale behaviour of microbial-induced calcium carbonate precipitation (micp). *Géotechnique*, 69(12):1086–1094, 01 2019. ISSN 0016-8505. doi: 10.1680/jgeot.18.P.031.
- 126 Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning: with Applications in R*. Springer: New York, NY, 2 edition, 2021. doi: 10.1007/978-1-0716-1418-1.
- 127 Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.
- 128 Aaron Fisher, Cynthia Rudin, and Francesca Dominici. All models are wrong, but many are useful: Learning a variable’s importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research*, 20(177): 1–81, 2019.

- 129 Son Bui Truong, Nu Nguyen Thi, and Duong Nguyen Thanh. An experimental study on unconfined compressive strength of soft soil-cement mixtures with or without ggbfs in the coastal area of vietnam. *Advances in Civil Engineering*, 2020(1):7243704, 2020.
- 130 Daniel Ribeiro, Raquel Néri, and Rafaela Cardoso. Influence of water content in the ucs of soil-cement mixtures for different cement dosages. *Procedia engineering*, 143: 59–66, 2016.
- 131 John P Malizia and Abdul Shakoor. Effect of water content and density on strength and deformation behavior of clay soils. *Engineering Geology*, 244:125–131, 2018.
- 132 Ho Anh Thu Nguyen, Duy Hoang Pham, Yonghan Ahn, Bee Lan Oo, and Benson Teck Heng Lim. Machine learning and sustainable geopolymer materials: A systematic review. *Materials Today Sustainability*, 30:101095, 2025. ISSN 2589-2347. doi: <https://doi.org/10.1016/j.mtsust.2025.101095>. URL <https://www.sciencedirect.com/science/article/pii/S2589234725000247>.
- 133 Bee Lan Oo, Anh Tuan Nguyen, Yonghan Ahn, and Benson Teck Heng Lim. Predicting the number of bidders in construction competitive bidding using explainable machine learning models. *Construction Innovation: Information, Process, Management*, 25(7):158–188, 2025. ISSN 1471-4175. doi: <https://doi.org/10.1108/CI-10-2024-0325>. URL <https://www.sciencedirect.com/science/article/pii/S1471417525000119>.
- 134 Henrique Echternacht, Luciana Campos, Alfeu Dias de Martinho, Danilo Pinto Moreira de Souza, Rodrigo Barbosa de Santis, Tiago Silveira Gontijo, Matteo Bodini, Angela Gorgoglione, Camila Martins Saporetti, and Leonardo Goliatt. Convolutional neural networks with transfer learning for natural river flow prediction in ungauged basins. *Scientific Reports*, 15(1):23873, 2025.
- 135 Qingliang Li, Ziyu Wang, Wei Shangguan, Lu Li, Yifei Yao, and Fanhua Yu. Improved daily smap satellite soil moisture prediction over china using deep learning model with transfer learning. *Journal of Hydrology*, 600:126698, 2021.

APÊNDICE A – Testes de normalidade e homogeneidade

Tabela 26 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D1.

Modelo	Métrica	Shapiro–Wilk (p)	Levene (p)	Premissas
AutoGluon	MAE	0,114	0,011	Violadas
AutoKeras	MAE	0,458	0,011	Violadas
FLAML	MAE	0,406	0,011	Violadas
H2O	MAE	0,027	0,011	Violadas
TPOT	MAE	0,003	0,011	Violadas
AutoGluon	MAPE	0,015	0,206	Violadas
AutoKeras	MAPE	0,152	0,206	Atendidas
FLAML	MAPE	0,013	0,206	Violadas
H2O	MAPE	0,174	0,206	Atendidas
TPOT	MAPE	0,003	0,206	Violadas
AutoGluon	R	0,033	0,222	Violadas
AutoKeras	R	0,378	0,222	Atendidas
FLAML	R	0,000	0,222	Violadas
H2O	R	0,085	0,222	Atendidas
TPOT	R	0,573	0,222	Atendidas
AutoGluon	R ²	0,196	0,012	Violadas
AutoKeras	R ²	0,902	0,012	Violadas
FLAML	R ²	0,001	0,012	Violadas
H2O	R ²	0,037	0,012	Violadas
TPOT	R ²	0,534	0,012	Violadas
AutoGluon	RMSE	0,050	0,489	Atendidas
AutoKeras	RMSE	0,386	0,489	Atendidas
FLAML	RMSE	0,022	0,489	Violadas
H2O	RMSE	0,167	0,489	Atendidas
TPOT	RMSE	0,359	0,489	Atendidas

Fonte: Elaborado pelo autor (2025).

Tabela 27 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D2.

Modelo	Métrica	Shapiro–Wilk (p)	Levene (p)	Premissas
AutoGluon	MAE	0,001	0,000	Violadas
AutoKeras	MAE	0,743	0,000	Violadas
FLAML	MAE	0,178	0,000	Violadas
H2O	MAE	0,001	0,000	Violadas
TPOT	MAE	0,587	0,000	Violadas
AutoGluon	MAPE	0,000	0,004	Violadas
AutoKeras	MAPE	0,157	0,004	Violadas
FLAML	MAPE	0,000	0,004	Violadas
H2O	MAPE	0,000	0,004	Violadas
TPOT	MAPE	0,032	0,004	Violadas
AutoGluon	R	0,003	0,017	Violadas
AutoKeras	R	0,002	0,017	Violadas
FLAML	R	0,002	0,017	Violadas
H2O	R	0,000	0,017	Violadas
TPOT	R	0,002	0,017	Violadas
AutoGluon	R ²	0,001	0,060	Violadas
AutoKeras	R ²	0,000	0,060	Violadas
FLAML	R ²	0,004	0,060	Violadas
H2O	R ²	0,000	0,060	Violadas
TPOT	R ²	0,000	0,060	Violadas
AutoGluon	RMSE	0,094	0,006	Violadas
AutoKeras	RMSE	0,371	0,006	Violadas
FLAML	RMSE	0,831	0,006	Violadas
H2O	RMSE	0,003	0,006	Violadas
TPOT	RMSE	0,435	0,006	Violadas

Fonte: Elaborado pelo autor (2025).

Tabela 28 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D3.

Modelo	Métrica	Shapiro–Wilk (p)	Levene (p)	Premissas
AutoGluon	MAE	0,122	0,000	Violadas
AutoKeras	MAE	0,803	0,000	Violadas
FLAML	MAE	0,936	0,000	Violadas
H2O	MAE	0,820	0,000	Violadas
TPOT	MAE	0,262	0,000	Violadas
AutoGluon	MAPE	0,006	0,000	Violadas
AutoKeras	MAPE	0,243	0,000	Violadas
FLAML	MAPE	0,563	0,000	Violadas
H2O	MAPE	0,210	0,000	Violadas
TPOT	MAPE	0,518	0,000	Violadas
AutoGluon	R	0,000	0,000	Violadas
AutoKeras	R	0,105	0,000	Violadas
FLAML	R	0,683	0,000	Violadas
H2O	R	0,001	0,000	Violadas
TPOT	R	0,010	0,000	Violadas
AutoGluon	R ²	0,000	0,000	Violadas
AutoKeras	R ²	0,159	0,000	Violadas
FLAML	R ²	0,529	0,000	Violadas
H2O	R ²	0,001	0,000	Violadas
TPOT	R ²	0,000	0,000	Violadas
AutoGluon	RMSE	0,023	0,000	Violadas
AutoKeras	RMSE	0,721	0,000	Violadas
FLAML	RMSE	0,879	0,000	Violadas
H2O	RMSE	0,269	0,000	Violadas
TPOT	RMSE	0,001	0,000	Violadas

Fonte: Elaborado pelo autor (2025).

Tabela 29 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D4.

Modelo	Métrica	Shapiro–Wilk (p)	Levene (p)	Premissas
AutoGluon	MAE	0,039	0,000	Violadas
AutoKeras	MAE	0,818	0,000	Violadas
FLAML	MAE	0,000	0,000	Violadas
H2O	MAE	0,001	0,000	Violadas
TPOT	MAE	0,025	0,000	Violadas
AutoGluon	MAPE	0,017	0,000	Violadas
AutoKeras	MAPE	0,220	0,000	Violadas
FLAML	MAPE	0,000	0,000	Violadas
H2O	MAPE	0,030	0,000	Violadas
TPOT	MAPE	0,009	0,000	Violadas
AutoGluon	R	0,051	0,000	Violadas
AutoKeras	R	0,939	0,000	Violadas
FLAML	R	0,000	0,000	Violadas
H2O	R	0,000	0,000	Violadas
TPOT	R	0,000	0,000	Violadas
AutoGluon	R ²	0,001	0,000	Violadas
AutoKeras	R ²	0,208	0,000	Violadas
FLAML	R ²	0,000	0,000	Violadas
H2O	R ²	0,000	0,000	Violadas
TPOT	R ²	0,001	0,000	Violadas
AutoGluon	RMSE	0,011	0,000	Violadas
AutoKeras	RMSE	0,673	0,000	Violadas
FLAML	RMSE	0,000	0,000	Violadas
H2O	RMSE	0,015	0,000	Violadas
TPOT	RMSE	0,004	0,000	Violadas

Fonte: Elaborado pelo autor (2025).

Tabela 30 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D5.

Modelo	Métrica	Shapiro–Wilk (p)	Levene (p)	Premissas
AutoGluon	MAE	0,260	0,103	Atendidas
AutoKeras	MAE	0,879	0,103	Atendidas
FLAML	MAE	0,523	0,103	Atendidas
H2O	MAE	0,887	0,103	Atendidas
TPOT	MAE	0,629	0,103	Atendidas
AutoGluon	MAPE	0,000	0,000	Violadas
AutoKeras	MAPE	0,100	0,000	Violadas
FLAML	MAPE	0,000	0,000	Violadas
H2O	MAPE	0,003	0,000	Violadas
TPOT	MAPE	0,000	0,000	Violadas
AutoGluon	R	0,000	0,051	Violadas
AutoKeras	R	0,151	0,051	Atendidas
FLAML	R	0,012	0,051	Violadas
H2O	R	0,003	0,051	Violadas
TPOT	R	0,014	0,051	Violadas
AutoGluon	R ²	0,002	0,123	Violadas
AutoKeras	R ²	0,044	0,123	Violadas
FLAML	R ²	0,003	0,123	Violadas
H2O	R ²	0,005	0,123	Violadas
TPOT	R ²	0,048	0,123	Violadas
AutoGluon	RMSE	0,412	0,640	Atendidas
AutoKeras	RMSE	0,694	0,640	Atendidas
FLAML	RMSE	0,179	0,640	Atendidas
H2O	RMSE	0,490	0,640	Atendidas
TPOT	RMSE	0,653	0,640	Atendidas

Fonte: Elaborado pelo autor (2025).

Tabela 31 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D6.

Modelo	Métrica	Shapiro–Wilk (p)	Levene (p)	Premissas
AutoGluon	MAE	0,286	0,000	Violadas
AutoKeras	MAE	0,321	0,000	Violadas
FLAML	MAE	0,920	0,000	Violadas
H2O	MAE	0,081	0,000	Violadas
TPOT	MAE	0,019	0,000	Violadas
AutoGluon	MAPE	0,663	0,247	Atendidas
AutoKeras	MAPE	0,373	0,247	Atendidas
FLAML	MAPE	0,374	0,247	Atendidas
H2O	MAPE	0,080	0,247	Atendidas
TPOT	MAPE	0,000	0,247	Violadas
AutoGluon	R	0,870	0,000	Violadas
AutoKeras	R	0,094	0,000	Violadas
FLAML	R	0,397	0,000	Violadas
H2O	R	0,780	0,000	Violadas
TPOT	R	0,339	0,000	Violadas
AutoGluon	R ²	0,696	0,000	Violadas
AutoKeras	R ²	0,216	0,000	Violadas
FLAML	R ²	0,368	0,000	Violadas
H2O	R ²	0,832	0,000	Violadas
TPOT	R ²	0,233	0,000	Violadas
AutoGluon	RMSE	0,941	0,024	Violadas
AutoKeras	RMSE	0,473	0,024	Violadas
FLAML	RMSE	0,945	0,024	Violadas
H2O	RMSE	0,300	0,024	Violadas
TPOT	RMSE	0,190	0,024	Violadas

Fonte: Elaborado pelo autor (2025).

Tabela 32 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D7.

Modelo	Métrica	Shapiro–Wilk (p)	Levene (p)	Premissas
AutoGluon	MAE	0,766	0,000	Violadas
AutoKeras	MAE	0,203	0,000	Violadas
FLAML	MAE	0,007	0,000	Violadas
H2O	MAE	0,000	0,000	Violadas
TPOT	MAE	0,266	0,000	Violadas
AutoGluon	MAPE	0,050	0,000	Violadas
AutoKeras	MAPE	0,001	0,000	Violadas
FLAML	MAPE	0,346	0,000	Violadas
H2O	MAPE	0,000	0,000	Violadas
TPOT	MAPE	0,088	0,000	Violadas
AutoGluon	R	0,000	0,000	Violadas
AutoKeras	R	0,206	0,000	Violadas
FLAML	R	0,001	0,000	Violadas
H2O	R	0,000	0,000	Violadas
TPOT	R	0,000	0,000	Violadas
AutoGluon	R ²	0,006	0,001	Violadas
AutoKeras	R ²	0,174	0,001	Violadas
FLAML	R ²	0,001	0,001	Violadas
H2O	R ²	0,000	0,001	Violadas
TPOT	R ²	0,001	0,001	Violadas
AutoGluon	RMSE	0,606	0,000	Violadas
AutoKeras	RMSE	0,171	0,000	Violadas
FLAML	RMSE	0,064	0,000	Violadas
H2O	RMSE	0,000	0,000	Violadas
TPOT	RMSE	0,241	0,000	Violadas

Fonte: Elaborado pelo autor (2025).

Tabela 33 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D8.

Modelo	Métrica	Shapiro–Wilk (p)	Levene (p)	Premissas
AutoGluon	MAE	0,271	0,000	Violadas
AutoKeras	MAE	0,624	0,000	Violadas
FLAML	MAE	0,543	0,000	Violadas
H2O	MAE	0,024	0,000	Violadas
TPOT	MAE	0,451	0,000	Violadas
AutoGluon	MAPE	0,609	0,001	Violadas
AutoKeras	MAPE	0,290	0,001	Violadas
FLAML	MAPE	0,634	0,001	Violadas
H2O	MAPE	0,004	0,001	Violadas
TPOT	MAPE	0,513	0,001	Violadas
AutoGluon	R	0,177	0,032	Violadas
AutoKeras	R	0,884	0,032	Violadas
FLAML	R	0,012	0,032	Violadas
H2O	R	0,000	0,032	Violadas
TPOT	R	0,231	0,032	Violadas
AutoGluon	R ²	0,119	0,001	Violadas
AutoKeras	R ²	0,918	0,001	Violadas
FLAML	R ²	0,013	0,001	Violadas
H2O	R ²	0,000	0,001	Violadas
TPOT	R ²	0,276	0,001	Violadas
AutoGluon	RMSE	0,618	0,000	Violadas
AutoKeras	RMSE	0,972	0,000	Violadas
FLAML	RMSE	0,308	0,000	Violadas
H2O	RMSE	0,061	0,000	Violadas
TPOT	RMSE	0,644	0,000	Violadas

Fonte: Elaborado pelo autor (2025).

Tabela 34 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D9.

Modelo	Métrica	Shapiro–Wilk (p)	Levene (p)	Premissas
AutoGluon	MAE	0,480	0,046	Violadas
AutoKeras	MAE	0,143	0,046	Violadas
FLAML	MAE	0,397	0,046	Violadas
H2O	MAE	0,000	0,046	Violadas
TPOT	MAE	0,768	0,046	Violadas
AutoGluon	MAPE	0,078	0,006	Violadas
AutoKeras	MAPE	0,047	0,006	Violadas
FLAML	MAPE	0,217	0,006	Violadas
H2O	MAPE	0,000	0,006	Violadas
TPOT	MAPE	0,003	0,006	Violadas
AutoGluon	R	0,188	0,113	Atendidas
AutoKeras	R	0,019	0,113	Violadas
FLAML	R	0,509	0,113	Atendidas
H2O	R	0,000	0,113	Violadas
TPOT	R	0,010	0,113	Violadas
AutoGluon	R ²	0,007	0,103	Violadas
AutoKeras	R ²	0,004	0,103	Violadas
FLAML	R ²	0,017	0,103	Violadas
H2O	R ²	0,001	0,103	Violadas
TPOT	R ²	0,013	0,103	Violadas
AutoGluon	RMSE	0,000	0,573	Violadas
AutoKeras	RMSE	0,215	0,573	Atendidas
FLAML	RMSE	0,000	0,573	Violadas
H2O	RMSE	0,008	0,573	Violadas
TPOT	RMSE	0,027	0,573	Violadas

Fonte: Elaborado pelo autor (2025).

Tabela 35 – Resultados dos testes de Shapiro–Wilk (normalidade) e Levene (homogeneidade) para cada métrica e sistema no conjunto D10.

Modelo	Métrica	Shapiro–Wilk (p)	Levene (p)	Premissas
AutoGluon	MAE	0,359	0,000	Violadas
AutoKeras	MAE	0,726	0,000	Violadas
FLAML	MAE	0,015	0,000	Violadas
H2O	MAE	0,000	0,000	Violadas
TPOT	MAE	0,423	0,000	Violadas
AutoGluon	MAPE	0,065	0,000	Violadas
AutoKeras	MAPE	0,108	0,000	Violadas
FLAML	MAPE	0,001	0,000	Violadas
H2O	MAPE	0,000	0,000	Violadas
TPOT	MAPE	0,000	0,000	Violadas
AutoGluon	R	0,051	0,000	Violadas
AutoKeras	R	0,024	0,000	Violadas
FLAML	R	0,979	0,000	Violadas
H2O	R	0,000	0,000	Violadas
TPOT	R	0,290	0,000	Violadas
AutoGluon	R ²	0,007	0,000	Violadas
AutoKeras	R ²	0,656	0,000	Violadas
FLAML	R ²	0,591	0,000	Violadas
H2O	R ²	0,000	0,000	Violadas
TPOT	R ²	0,655	0,000	Violadas
AutoGluon	RMSE	0,016	0,000	Violadas
AutoKeras	RMSE	0,622	0,000	Violadas
FLAML	RMSE	0,026	0,000	Violadas
H2O	RMSE	0,000	0,000	Violadas
TPOT	RMSE	0,371	0,000	Violadas

Fonte: Elaborado pelo autor (2025).