v7.n12.2026

REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218

METODOLOGIA DE UTILIZAÇÃO DO *FRAMEWORK* OPTUNA PARA OTIMIZAÇÃO DE HIPERPARÂMETROS DE MODELO DE APRENDIZADO DE MÁQUINA DE REGRESSÃO

METHODOLOGY FOR USING THE OPTUNA FRAMEWORK TO OPTIMIZE HYPERPARAMETERS OF A REGRESSION MACHINE LEARNING MODEL

METODOLOGÍA DE USO DEL FRAMEWORK OPTUNA PARA OPTIMIZAR HIPERPARÁMETROS DE UN MODELO DE APRENDIZAJE AUTOMÁTICO DE REGRESIÓN

Fabricio Ribeiro Seppe[1]

https://doi.org/10.47820/recima21.v7i12.9477

PUBLICADO: 12/2026

RESUMO

Este estudo apresenta uma metodologia de utilização do *framework* Optuna para otimizar hiperparâmetros de modelos de aprendizado de máquina de regressão, ilustrada por sua aplicação ao XGBoost. O objetivo foi organizar a análise dos registros de otimização para apoiar a definição de novos intervalos de busca. Foram utilizadas onze bases públicas do UC Irvine Machine Learning Repository, desdobradas em quinze tarefas, com preparação dos dados, codificação de variáveis categóricas e normalização. O procedimento, denominado MOHOC, combina os registros das tentativas do Optuna com visualizações em Microsoft Excel, permitindo ao analista examinar concentrações de resultados e redefinir os intervalos dos hiperparâmetros. Foram realizados quatro ciclos de otimização por tarefa, com 1500 tentativas no primeiro ciclo. Ao final, a raiz do erro quadrático médio (RMSE) apresentou redução média de 15,91% em relação às configurações padrão e de 4,24% em relação ao primeiro ciclo. Os resultados descrevem a evolução observada durante a busca, sem constituir uma avaliação independente de generalização ou isolar o efeito do refinamento manual do aumento do orçamento de otimização. A contribuição consiste na organização de um procedimento prático de análise e condução dos ciclos de busca.

PALAVRAS-CHAVE: Aprendizado de máquina. Regressão. Otimização de hiperparâmetros. Optuna. XGBoost.

ABSTRACT

This study presents a methodology for using the Optuna framework to optimize hyperparameters in regression machine learning models, illustrated by its application to XGBoost. The objective was to organize the analysis of optimization logs to support the definition of new search intervals. Eleven public datasets from the UC Irvine Machine Learning Repository were used to define fifteen tasks, with data preparation, categorical encoding, and normalization. The procedure, named MOHOC, combines Optuna trial logs with Microsoft Excel visualizations, enabling the analyst to examine concentrations of results and redefine hyperparameter intervals. Four optimization cycles were performed for each task, with 1500 trials in the first cycle. At the end, the root mean squared error (RMSE) decreased by an average of 15.91% relative to the default configurations and 4.24% relative to the first cycle. These results describe the observed search progression; they do not provide an independent assessment of generalization or isolate the effect of manual refinement from the increased optimization budget. The contribution is a practical procedure for analyzing and guiding successive search cycles.

KEYWORDS: *Machine learning. Regression. Hyperparameter optimization. Optuna. XGBoost.*

[1] Escola Superior de Propaganda e Marketing (ESPM). Engenheiro civil (UEA) especialista em Data Science e Analytics e mestrando na ESPM.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

REVISTA CIENTÍFICA - RECIMA21 ISSN 2675-6218

RESUMEN
Este estudio presenta una metodología de uso del framework Optuna para optimizar hiperparámetros de modelos de aprendizaje automático para regresión, ilustrada mediante su aplicación a XGBoost. El objetivo fue organizar el análisis de los registros de optimización para apoyar la definición de nuevos intervalos de búsqueda. Se utilizaron once conjuntos de datos públicos del UC Irvine Machine Learning Repository para definir quince tareas, con preparación de datos, codificación de variables categóricas y normalización. El procedimiento, denominado MOHOC, combina los registros de los ensayos de Optuna con visualizaciones en Microsoft Excel, permitiendo al analista examinar concentraciones de resultados y redefinir los intervalos de los hiperparámetros. Se realizaron cuatro ciclos de optimización por tarea, con 1500 ensayos en el primero. Al finalizar, la raíz del error cuadrático medio (RMSE) presentó una reducción media del 15,91% respecto de las configuraciones predeterminadas y del 4,24% respecto del primer ciclo. Los resultados describen la evolución observada durante la búsqueda, sin constituir una evaluación independiente de generalización ni aislar el efecto del refinamiento manual del incremento del presupuesto de optimización. La contribución consiste en un procedimiento práctico para analizar y orientar los ciclos de búsqueda.

PALABRAS CLAVE*: Aprendizaje automático. Regresión. Optimización de hiperparámetros. Optuna. XGBoost.*

1. INTRODUÇÃO

O ajuste de hiperparâmetros é uma etapa relevante na aplicação de modelos de aprendizado de máquina, pois configurações distintas podem produzir desempenhos diferentes para um mesmo conjunto de dados. A definição dos intervalos de busca também condiciona as configurações que poderão ser avaliadas. Assim, além da escolha do algoritmo de otimização, é necessário examinar como os resultados intermediários podem orientar a continuidade da busca (Yang; Shami, 2020).

Aplicações em previsão de carga e consumo de energia ilustram a utilização de modelos preditivos e de estratégias de combinação de modelos em diferentes contextos (Alquthami *et al.*, 2022; Al-Rakhami et al., 2019; Tan et al., 2020). Neste trabalho, o foco recai sobre o XGBoost, que implementa uma abordagem de árvores de decisão impulsionadas por gradiente (Chen; Guestrin, 2016), e sobre o Optuna, uma ferramenta de otimização de hiperparâmetros (Akiba et al., 2019).

O problema investigado é como organizar a inspeção dos resultados das tentativas do Optuna para apoiar a redefinição manual dos intervalos de hiperparâmetros em ciclos sucessivos. A justificativa é de natureza prática: uma busca inicial pode não explorar regiões que apresentem resultados favoráveis, e a consulta isolada ao melhor resultado não revela toda a distribuição das tentativas realizadas.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

O objetivo geral foi desenvolver e aplicar um procedimento de análise visual dos registros do Optuna para orientar o refinamento manual do espaço de busca em modelos de regressão, utilizando o XGBoost como aplicação experimental. Os objetivos específicos foram preparar conjuntos de dados públicos; registrar desempenhos com as configurações padrão; organizar os resultados das tentativas em uma ferramenta de visualização; conduzir quatro ciclos de busca; e descrever as variações observadas na métrica de avaliação.

A contribuição proposta é um fluxo de trabalho que articula otimização automatizada e interpretação humana por meio de planilhas e mapas de calor. O estudo tem caráter aplicado e exploratório: examina a utilização do procedimento em quinze tarefas derivadas de onze bases públicas, sem propor um novo algoritmo de otimização ou demonstrar superioridade geral sobre outras estratégias de busca.

2. REFERENCIAL TEÓRICO

2.1. Preparação dos dados e regressão

A preparação dos dados, também denominada *data wrangling*, envolve organizar, transformar e tratar informações para sua utilização analítica (Boehmke, 2016). Conforme o conjunto de dados, pode incluir conversão de tipos, tratamento de valores ausentes, codificação de variáveis categóricas e normalização. Essas operações devem ser descritas de modo a permitir a interpretação dos resultados e a identificação das limitações do procedimento.

Modelos de aprendizado de máquina identificam padrões em dados de treinamento para realizar tarefas como classificação, regressão e agrupamento. Em regressão, estimam valores numéricos para novas observações. Seu desempenho depende das características dos dados e das condições de treinamento e avaliação.

O XGBoost (*eXtreme Gradient Boosting*), utilizado aqui para regressão, combina árvores de decisão por impulsionamento por gradiente. A previsão resulta da soma das contribuições das árvores. O treinamento utiliza derivadas de primeira e segunda ordem da função de perda e termos de regularização para controlar a complexidade do modelo (Zhang; Gong, 2020; Chen; Guestrin, 2016).

2.2. Otimização de hiperparâmetros e Optuna

Os parâmetros ajustados durante o treinamento diferem dos hiperparâmetros definidos para configurar o modelo e seu processo de aprendizado. Na otimização de hiperparâmetros, o *Grid Search* percorre uma grade de configurações, enquanto o *Random Search* amostra

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

configurações de um espaço definido. O custo dessas estratégias depende do tamanho do espaço e do orçamento de avaliação disponível (Yang; Shami, 2020; Demidova; Filatov, 2022). Comparações entre estratégias devem considerar o problema, a métrica e as condições de execução (Arden; Safitri, 2022).

O Optuna é um *framework* de otimização que permite definir uma função objetivo e explorar diferentes configurações de hiperparâmetros (Akiba *et al.*, 2019). Sua arquitetura admite diferentes estratégias de amostragem e mecanismos de interrupção antecipada de tentativas. A disponibilidade desses recursos não implica que todos sejam utilizados em uma aplicação específica. Neste estudo, o procedimento apresentado concentra-se na análise dos registros das tentativas e na redefinição dos intervalos entre estudos.

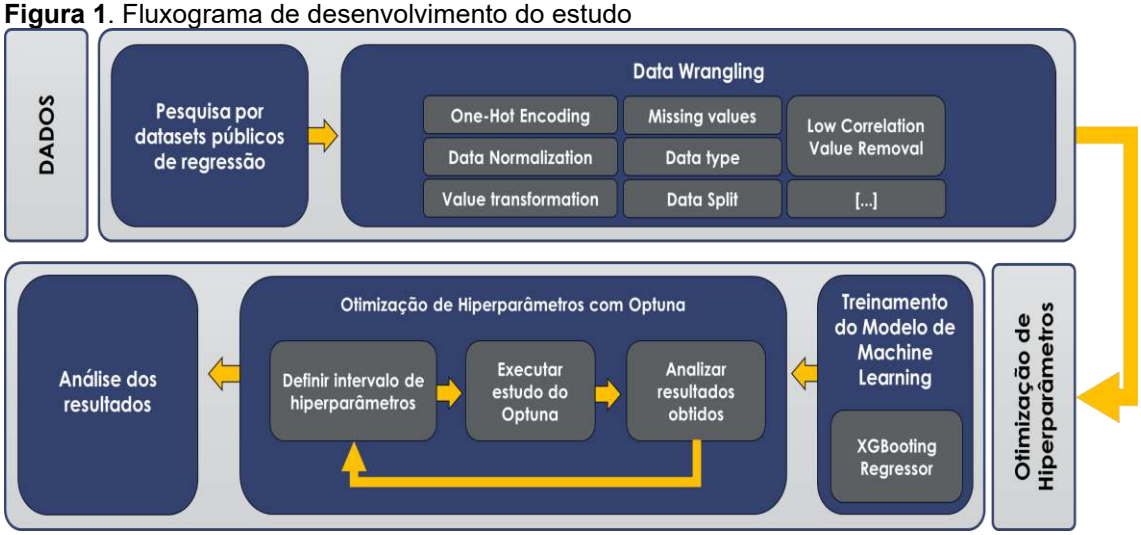
Uma tentativa, ou *trial*, corresponde à avaliação de uma configuração de hiperparâmetros. Um estudo reúne tentativas orientadas por uma função objetivo. Além de seu uso em modelos de árvores, a otimização de configurações integra aplicações de modelos híbridos de previsão, como a previsão de radiação ultravioleta estudada por Prasad et al. (2022).

3. METODOLOGIA

3.1. Delineamento e conjuntos de dados

Foi desenvolvido um estudo computacional aplicado e exploratório, composto pela preparação dos dados, treinamento de modelos de referência, otimização com Optuna e análise visual das tentativas para definição de ciclos subsequentes. A Figura 1 apresenta o fluxo geral de desenvolvimento.

Figura 1. Fluxograma de desenvolvimento do estudo



Fonte: elaboração própria (2026).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Foram selecionadas onze bases públicas do UC Irvine Machine Learning Repository para aplicação do procedimento. A Tabela 1 apresenta suas características.

Tabela 1. Bases públicas selecionadas no UC Irvine Machine Learning Repository

ID	*Dataset*	*Área*	*DOI*	*Instâncias*	*Atributos*
01	*Communities and Crime*	*Social Science*	10.24432/C53W3X	1994	127
02	*Productivity Prediction of Garment Employees*	*Business*	10.24432/C51S6D	1197	14
03	*Combined Cycle Power Plant*	*Computer Science*	10.24432/C5002N	9568	4
04	*Solar Flare*	*Physics and Chemistry*	10.24432/C5530G	1389	10
05	*Parkinsons Telemonitoring*	*Health and Medicine*	10.24432/C5ZS3N	5875	19
06	*Metro Interstate Traffic Volume*	*Public Transportation*	10.24432/C5X60B	48204	8
07	*Steel Industry Energy Consumption*	*Physics and Chemistry*	10.24432/C52G8C	35040	9
08	*Accelerometer*	*Physics and Chemistry*	10.24432/C5Q61V	153000	5
09	*AI4I 2020 Predictive Maintenance Dataset*	*Computer Science*	10.24432/C5HS5C	10000	6
10	*Seoul Bike Sharing Demand*	*Business*	10.24432/C5F62R	8760	13
11	*Bike Sharing*	*Social Science*	10.24432/C5W894	17389	13

Fonte: elaboração própria (2026).

As onze bases da Tabela 1 originaram quinze tarefas de regressão. *Solar Flare* foi utilizada com três variáveis-alvo; *Parkinsons Telemonitoring*, com duas; e *Bike Sharing*, em suas versões diária e horária. As demais bases deram origem a uma tarefa cada. Portanto, as quinze tarefas não representam quinze fontes de dados independentes.

3.2. Preparação e particionamento dos dados

Os códigos foram desenvolvidos em Python, no ambiente Anaconda, utilizando JupyterLab. A implementação utiliza bibliotecas de análise de dados e aprendizado de máquina, um dicionário de configuração de cada conjunto de dados e uma função de inspeção das colunas. O dicionário armazena a identificação da variável-alvo e as opções de tratamento aplicáveis a cada tarefa.

A preparação incluiu tentativas de conversão de tipos para *float*, *int* e *datetime*; tratamento de valores ausentes com *pandas.DataFrame.dropna*; definição da coluna de índice; e codificação de variáveis categóricas por *One-Hot Encoding*, conforme a configuração de cada conjunto de dados.

A matriz de correlação foi calculada com pandas e representada com seaborn.*heatmap*. Em seguida, foi aplicado um filtro de variáveis baseado no valor absoluto da correlação com o alvo,

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

segundo o limiar definido no dicionário de tratamento. Esse filtro representa uma escolha operacional do estudo: correlação linear baixa não demonstra ausência de contribuição preditiva, especialmente em modelos capazes de representar relações não lineares.

Após a seleção das variáveis, os dados foram particionados com *train_test_split*, com random_state=1. A normalização min-max dos preditores e da variável-alvo foi ajustada com os dados de treino e aplicada à outra partição por transformação. Como a variável-alvo foi normalizada, a escala dos erros apresentados depende dessa transformação.

A segunda partição é denominada teste nos scripts, mas foi utilizada repetidamente no cálculo da função objetivo durante a otimização. Ela exerce, portanto, a função de validação para seleção de hiperparâmetros. O protocolo apresentado não inclui uma avaliação final em uma terceira partição independente. Os resultados devem ser interpretados como desempenho observado durante o processo de busca.

3.3. Modelo de referência e métrica de avaliação

Para cada tarefa foi treinado um modelo *XGBoost Regressor* com os hiperparâmetros padrão utilizados na implementação original. Os desempenhos de referência foram registrados para comparação com os resultados dos quatro ciclos de otimização. A Tabela 2 apresenta esses valores.

Tabela 2. Desempenho de referência por tarefa

ID	*Dataset*	*RMSE*
01	*Communities and Crime*	0,1455097594
02	*Productivity Prediction of Garment Employees*	0,0910963492
03	*Combined Cycle Power Plant*	0,0412793049
04	*Solar Flare (Common Flares)*	0,1035761657
05	*Solar Flare (Moderate Flares)*	0,0840333003
06	*Solar Flare (Severe Flares)*	0,0483745421
07	*Parkinsons Telemonitoring (Total)*	0,1506086993
08	*Parkinsons Telemonitoring (Motor)*	0,1623243167
09	*Metro Interstate Traffic Volume*	0,1297878603
10	*Steel Industry Energy Consumption*	0,0070842752
11	*Accelerometer*	0,0386591962
12	*Predictive Maintenance*	0,0366795586
13	*Seoul Bike Sharing Demand*	0,0613538420
14	*Bike Sharing (Day)*	0,0132820243
15	*Bike Sharing (Hour)*	0,0046946870

Fonte: elaboração própria (2026).

A função objetivo do Optuna minimizou o erro quadrático médio (MSE), calculado com *mean_squared_error*. Após a exportação dos registros, foi calculada no Microsoft Excel a raiz quadrada de cada valor de MSE para obter o RMSE apresentado nas tabelas e nos gráficos. Para N observações, valores observados O_i e valores previstos P_i, sua definição é:

$$RMSE = \sqrt{[(1/N) \sum_{i=1}^{N} (O_i - P_i)^2]}$$

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Valores menores indicam menor erro na partição utilizada para avaliação. Como a raiz quadrada é estritamente crescente para valores não negativos, a conversão de MSE para RMSE preserva a ordenação das tentativas. As reduções percentuais apresentadas referem-se aos valores de RMSE, e as tabelas mantêm a precisão numérica registrada nos experimentos.

3.4. Ciclos de otimização e ferramenta MOHOC

A função objective(*trial*) instancia e treina o XGBoost, calcula previsões na partição de avaliação e devolve o erro ao Optuna. Foram definidos intervalos iniciais comuns às tarefas, posteriormente ajustados conforme os resultados. Por exemplo, *max_depth* foi explorado entre 2 e 30, e *learning_rate* entre 0,03 e 1.

Após cada estudo, os registros foram organizados em um *dataframe* denominado *Total Logs* e exportados em formato CSV, incluindo a tarefa, a métrica e os hiperparâmetros de cada tentativa. Esses registros alimentam a ferramenta MOHOC (*Manual Optuna Hyperparameter Optimization Cycle*), destinada à análise manual dos ciclos de otimização de hiperparâmetros.

A versão 1.0.0 da MOHOC está disponível em acesso aberto no Zenodo (Seppe, 2026), sob licença MIT, em https://doi.org/10.5281/zenodo.24096106. O depósito inclui a planilha-modelo para preenchimento, uma planilha com exemplo ilustrativo do Cubist e guias de uso em português e inglês. O exemplo Cubist destina-se à demonstração da ferramenta e não corresponde aos experimentos com XGBoost apresentados neste artigo.

Os registros foram importados para o Microsoft Excel. Para cada hiperparâmetro, definiram-se os limites do intervalo e sua divisão em n subintervalos. As tentativas foram ordenadas pelo erro e receberam classificação normalizada entre 0 e 1, sendo 1 a melhor posição. A linha *Class med* apresenta a média dessas classificações em cada subintervalo. Os intervalos incluem o limite inferior e excluem o superior, exceto o último, que inclui ambos (Figura 2).

A linha *Class analysis* compara essa média com o limiar *Class cutoff*: *Keep* sugere manter o subintervalo quando a média é igual ou superior ao corte; *Discard* sugere descartá-lo quando é inferior. A regra utiliza a classificação normalizada, e não o valor bruto do RMSE. A decisão final permanece a cargo do analista.

A incidência relativa é calculada como m/(T/k), em que m é a contagem no subintervalo, T é o total contabilizado pela planilha e k é o número efetivo de subintervalos; com *Split last* ativado, k = n + 1. A referência T/k corresponde à distribuição uniforme. Em *Relative inc class*, Acc (*Acceptable*) indica incidência relativa igual ou superior ao *Incidence cutoff*; *Low* indica valor inferior. Trata-se de um critério operacional de cobertura amostral, não de uma garantia de precisão estatística.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Figura 2. Exemplo da interface MOHOC: distribuição por subintervalo e indicadores de classificação e incidência

Atb	
neighbors	
Range start	2
Range limit	9
Subinterval splits	7
Split last	Yes
Class cutoff	0,49
Incidence cutoff	20%
Increments	1
Atb Column Name	Data_Ordering[r

ID		1	2	3	4
Subinterval	>=	2	3	4	5
	<	3	4	5	6
Class med		0,468	0,483	0,464	0,513
Absolute class dif		-0,022	-0,007	-0,026	0,023
relative class dif		95,4%	98,5%	94,6%	104,7%
Above class Incidencies		45,7%	48,2%	45,7%	52,5%
Below class incidencies		54,3%	51,8%	54,3%	47,5%
Incidence		739	770	589	1630
Global inc		24,2%	25,3%	19,3%	53,5%
Relative inc		193,9%	202,0%	154,5%	427,7%
Relativ Inc class		Acc	Acc	Acc	Acc
Class analysis		Discard	Discard	Discard	Keep

Fonte: elaboração própria (2026).

Para cada subintervalo do hiperparâmetro em análise, calculam-se, separadamente para cada um dos demais hiperparâmetros, as médias de seus valores em dois grupos de tentativas. O CPM (*Cutoff Pass Mean*) corresponde à média no grupo com classificação normalizada igual ou superior ao corte definido para o conjunto de tentativas em *Class cutoff*; o CFM (*Cutoff Fail Mean*), à média no grupo abaixo desse corte. Em ambos os cálculos, consideram-se apenas as tentativas pertencentes ao subintervalo analisado. Essas médias apoiam a inspeção das combinações entre hiperparâmetros e desempenho, sem demonstrar relações causais.

Figura 3. Médias dos demais hiperparâmetros por subintervalo do hiperparâmetro analisado

n_rules	CPM	200,8284	206,0081	206,1599	213,2632
	CFM	188,5885	195,9073	189,7188	199,8968
n_committees	CPM	77,7219	77,5121	79,4610	81,1357
	CFM	74,1471	75,6917	78,2406	78,7574
neighbors	CPM	2	3	4	5
	CFM	2	3	4	5
extrapolation	CPM	0,27	0,27	0,27	0,27
	CFM	0,22	0,21	0,21	0,20
random_state	CPM	0,0000	0,0000	0,0000	0,0000
	CFM	0,0000	0,0000	0,0000	0,0000
auto	CPM	0,0237	0,0243	0,0409	0,0292
	CFM	0,1122	0,0827	0,1188	0,1032
sample	CPM	0,0000	0,0000	0,0000	0,0000
	CFM	0,0000	0,0000	0,0000	0,0000

Fonte: elaboração própria (2026).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Os mapas de calor apresentam os subintervalos do hiperparâmetro no eixo horizontal e cinquenta faixas de classificação no eixo vertical. A classificação varia de 0% a 100%, com agrupamento de 2% em 2%; valores maiores representam melhores posições. Cada célula contém a contagem de tentativas na combinação entre faixa e subintervalo. Tons de azul mais escuros representam maior contagem (Figura 4).

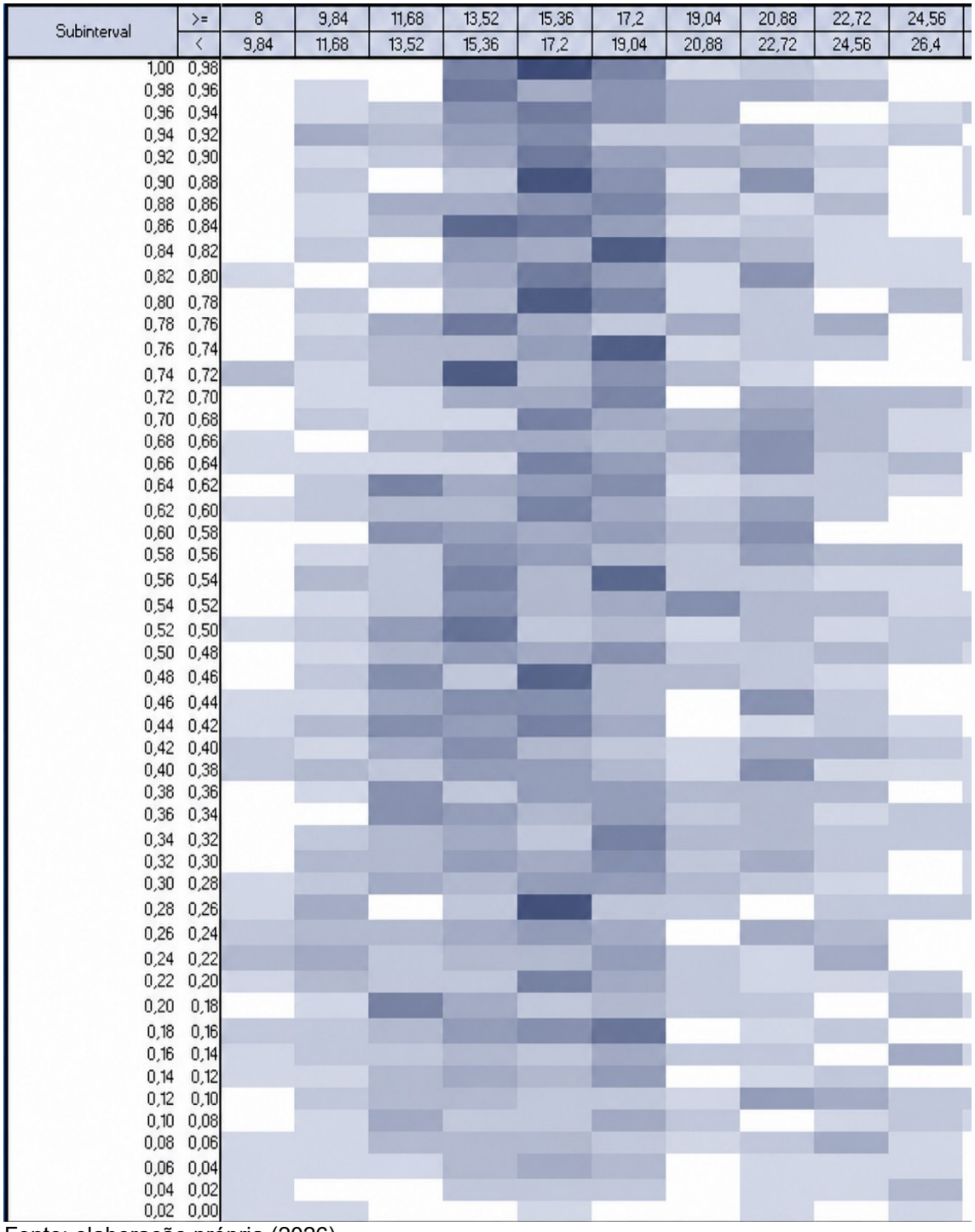
Figura 4. Painel A. Visão geral da distribuição das tentativas por subintervalo e faixa de classificação

13,52	15,36	17,2
15,36	17,2	19,04
7	12	7
8	4	6
6	8	6
5	7	2
4	8	5
2	11	6
4	6	7
9	8	5
5	4	10
4	8	5
3	10	7
8	4	2
3	5	10
10	3	6

Fonte: elaboração própria (2026).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Figura 4. Painel B. Detalhe do mapa de calor: subintervalos nas colunas e faixas de classificação nas linhas



Fonte: elaboração própria (2026).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

A definição dos novos intervalos foi realizada manualmente pelo analista, mediante interpretação conjunta das visualizações. O procedimento depende de decisões sobre a divisão em subintervalos, o limiar de classificação e a amplitude da busca subsequente. Essas decisões não constituem uma regra automatizada de atualização.

Foram realizados quatro ciclos de otimização por tarefa. O primeiro utilizou intervalos comuns e 1500 tentativas; os três seguintes utilizaram intervalos definidos com apoio da MOHOC. As comparações descritivas consideraram o desempenho padrão, o melhor resultado ao final do primeiro ciclo e o melhor resultado ao final dos quatro ciclos.

4. RESULTADOS E DISCUSSÃO

4.1. Análise visual e redefinição dos intervalos

Em um dos exemplos, *learning_rate* apresentou forte concentração de tentativas (incidência 700% superior à média de distribuição por intervalo) e classificações favoráveis entre 0,139224 e 0,1607152. O intervalo de busca foi restringido a essa faixa no ciclo subsequente.

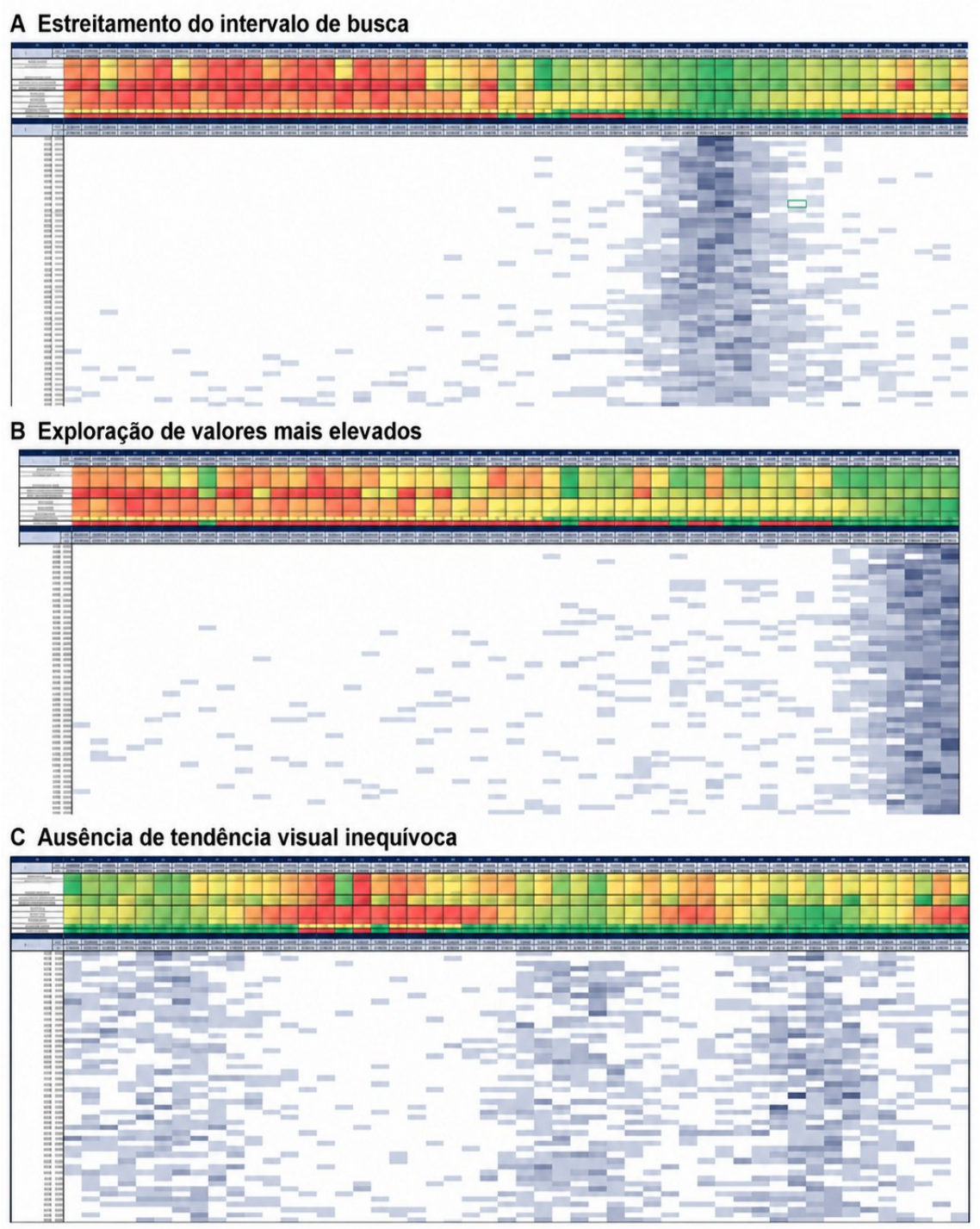
Outras análises indicaram concentrações de resultados favoráveis próximas aos limites do intervalo vigente. Nesses casos, os intervalos dos ciclos seguintes incluíram a região favorável já explorada e estenderam os limites em fatores de 2 a 10.

Quando o padrão visual foi menos claro, a análise considerou conjuntamente a incidência relativa, as médias condicionais dos demais hiperparâmetros e os intervalos definidos para eles.

Os exemplos ilustram concentração em uma faixa interna, concentração próxima aos limites e ausência de tendência visual inequívoca. A Figura 5 apresenta visões gerais desses padrões; a Figura 6 permite examinar os valores numéricos de subintervalos ilustrativos. As letras A, B e C identificam esses três tipos de situação. A interpretação conjunta dos indicadores apoia a decisão do analista, considerando as interações entre hiperparâmetros.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

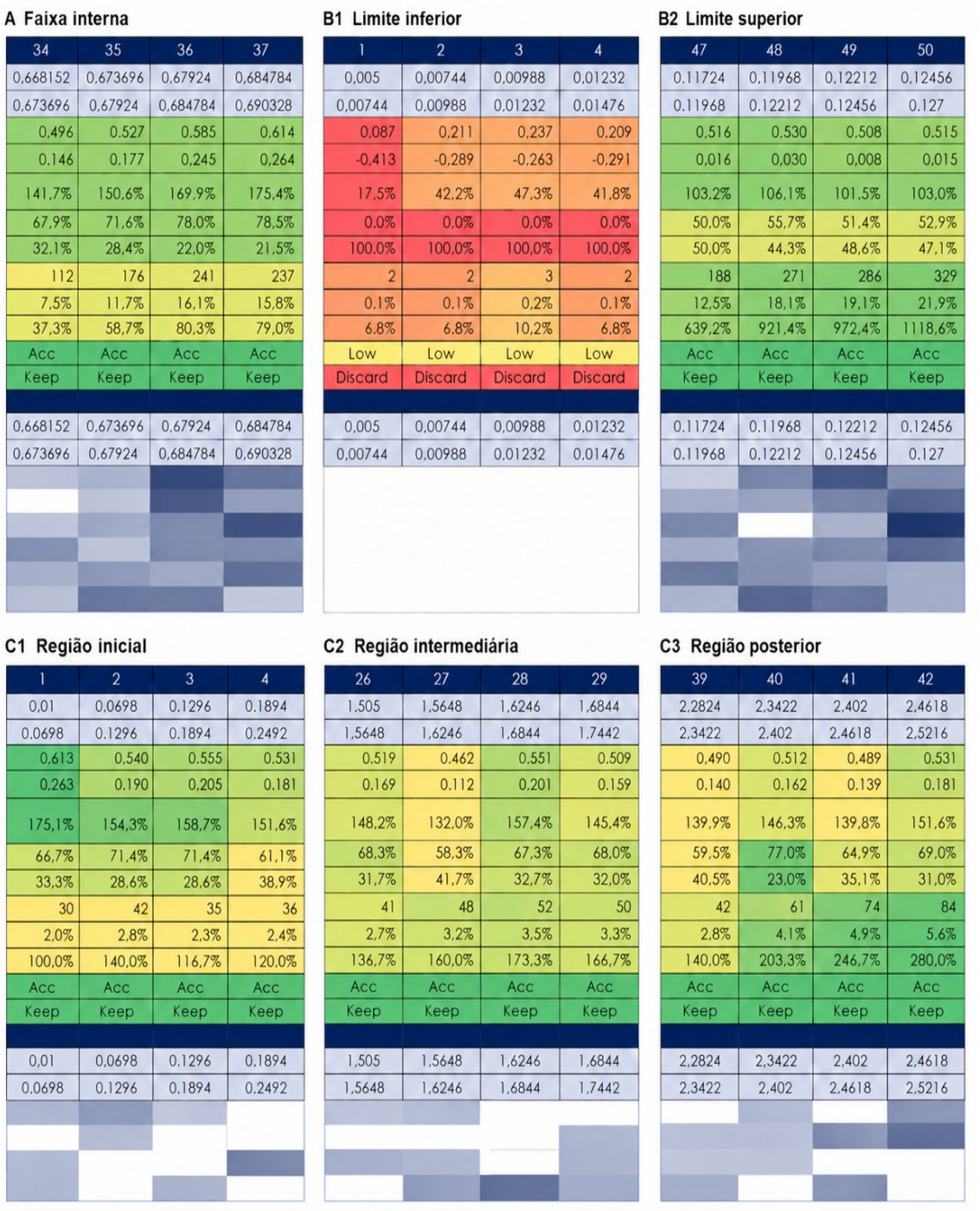
Figura 5. Exemplos de resultados dos heatmaps nos intervalos de busca — A: estreitamento; B: exploração de valores mais elevados; C: ausência de tendência visual inequívoca



Fonte: elaboração própria (2026).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Figura 6. Detalhes numéricos dos tipos de situação ilustrados na Figura 5 — A: faixa interna; B1 e B2: limites inferior e superior; C1, C2 e C3: regiões sem tendência inequívoca



A Faixa interna

34	35	36	37
0,668152	0,673696	0,67924	0,684784
0,673696	0,67924	0,684784	0,690328
0,496	0,527	0,585	0,614
0,146	0,177	0,245	0,264
141,7%	150,6%	169,9%	175,4%
67,9%	71,6%	78,0%	78,5%
32,1%	28,4%	22,0%	21,5%
112	176	241	237
7,5%	11,7%	16,1%	15,8%
37,3%	58,7%	80,3%	79,0%
Acc	Acc	Acc	Acc
Keep	Keep	Keep	Keep
0,668152	0,673696	0,67924	0,684784
0,673696	0,67924	0,684784	0,690328

B1 Limite inferior

1	2	3	4
0,005	0,00744	0,00988	0,01232
0,00744	0,00988	0,01232	0,01476
0,087	0,211	0,237	0,209
-0,413	-0,289	-0,263	-0,291
17,5%	42,2%	47,3%	41,8%
0,0%	0,0%	0,0%	0,0%
100,0%	100,0%	100,0%	100,0%
2	2	3	2
0,1%	0,1%	0,2%	0,1%
6,8%	6,8%	10,2%	6,8%
Low	Low	Low	Low
Discard	Discard	Discard	Discard
0,005	0,00744	0,00988	0,01232
0,00744	0,00988	0,01232	0,01476

B2 Limite superior

47	48	49	50
0.11724	0,11968	0,12212	0,12456
0.11968	0,12212	0,12456	0.127
0,516	0,530	0,508	0,515
0,016	0,030	0,008	0,015
103,2%	106,1%	101,5%	103,0%
50,0%	55,7%	51,4%	52,9%
50,0%	44,3%	48,6%	47,1%
188	271	286	329
12,5%	18,1%	19,1%	21,9%
639,2%	921,4%	972,4%	1118,6%
Acc	Acc	Acc	Acc
Keep	Keep	Keep	Keep
0.11724	0,11968	0,12212	0,12456
0.11968	0,12212	0,12456	0.127

C1 Região inicial

1	2	3	4
0,01	0,0698	0,1296	0,1894
0,0698	0,1296	0,1894	0,2492
0,613	0,540	0,555	0,531
0,263	0,190	0,205	0,181
175,1%	154,3%	158,7%	151,6%
66,7%	71,4%	71,4%	61,1%
33,3%	28,6%	28,6%	38,9%
30	42	35	36
2,0%	2,8%	2,3%	2,4%
100,0%	140,0%	116,7%	120,0%
Acc	Acc	Acc	Acc
Keep	Keep	Keep	Keep
0,01	0,0698	0,1296	0,1894
0,0698	0,1296	0,1894	0,2492

C2 Região intermediária

26	27	28	29
1,505	1,5648	1,6246	1,6844
1,5648	1,6246	1,6844	1,7442
0,519	0,462	0,551	0,509
0,169	0,112	0,201	0,159
148,2%	132,0%	157,4%	145,4%
68,3%	58,3%	67,3%	68,0%
31,7%	41,7%	32,7%	32,0%
41	48	52	50
2,7%	3,2%	3,5%	3,3%
136,7%	160,0%	173,3%	166,7%
Acc	Acc	Acc	Acc
Keep	Keep	Keep	Keep
1,505	1,5648	1,6246	1,6844
1,5648	1,6246	1,6844	1,7442

C3 Região posterior

39	40	41	42
2,2824	2,3422	2,402	2,4618
2,3422	2,402	2,4618	2,5216
0,490	0,512	0,489	0,531
0,140	0,162	0,139	0,181
139,9%	146,3%	139,8%	151,6%
59,5%	77,0%	64,9%	69,0%
40,5%	23,0%	35,1%	31,0%
42	61	74	84
2,8%	4,1%	4,9%	5,6%
140,0%	203,3%	246,7%	280,0%
Acc	Acc	Acc	Acc
Keep	Keep	Keep	Keep
2,2824	2,3422	2,402	2,4618
2,3422	2,402	2,4618	2,5216

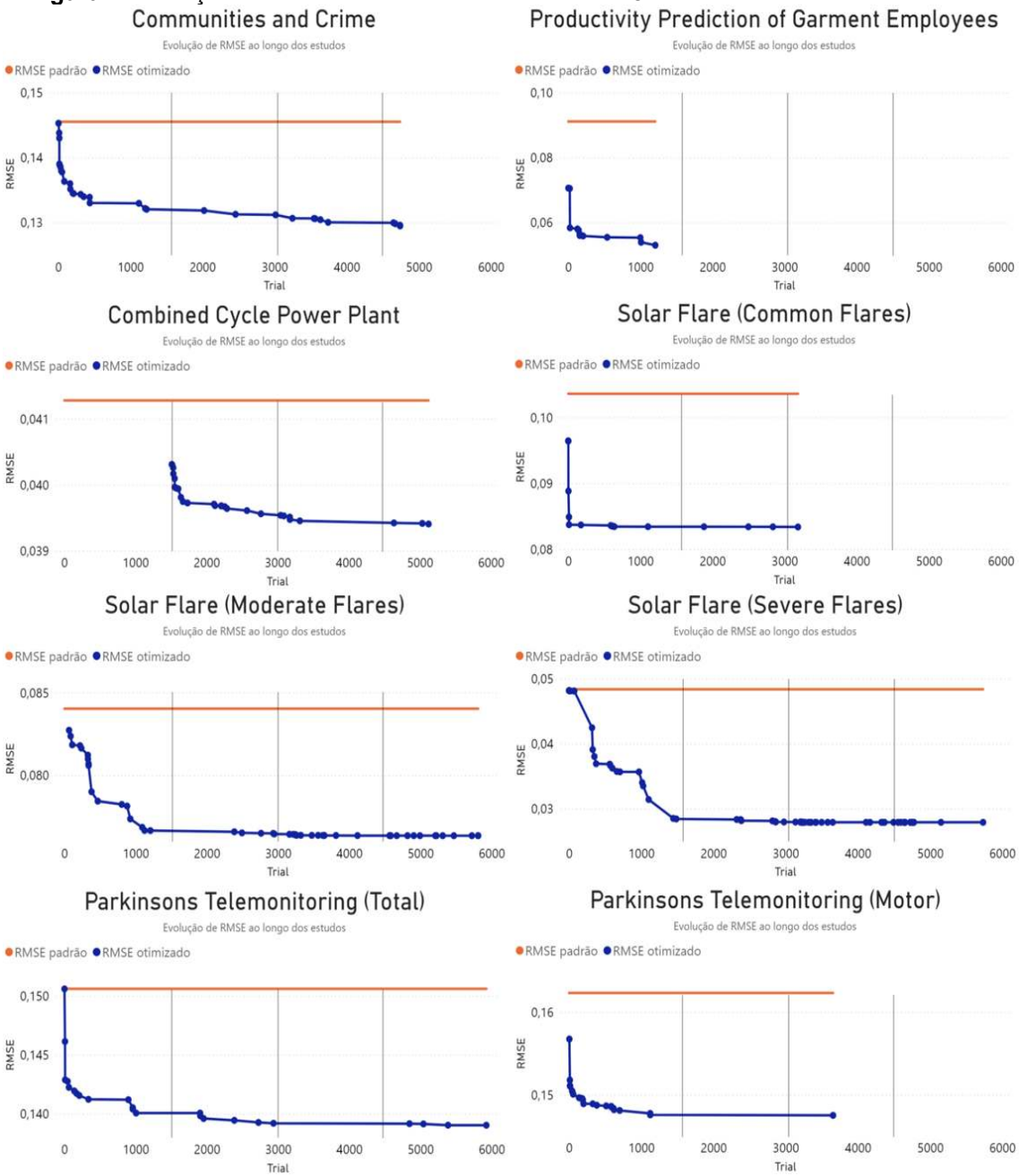
Fonte: elaboração própria (2026).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

4.2. Evolução dos resultados ao longo dos ciclos

As Figuras 7 e 8 apresentam a evolução do RMSE ao longo dos quatro ciclos de otimização nas quinze tarefas de regressão.

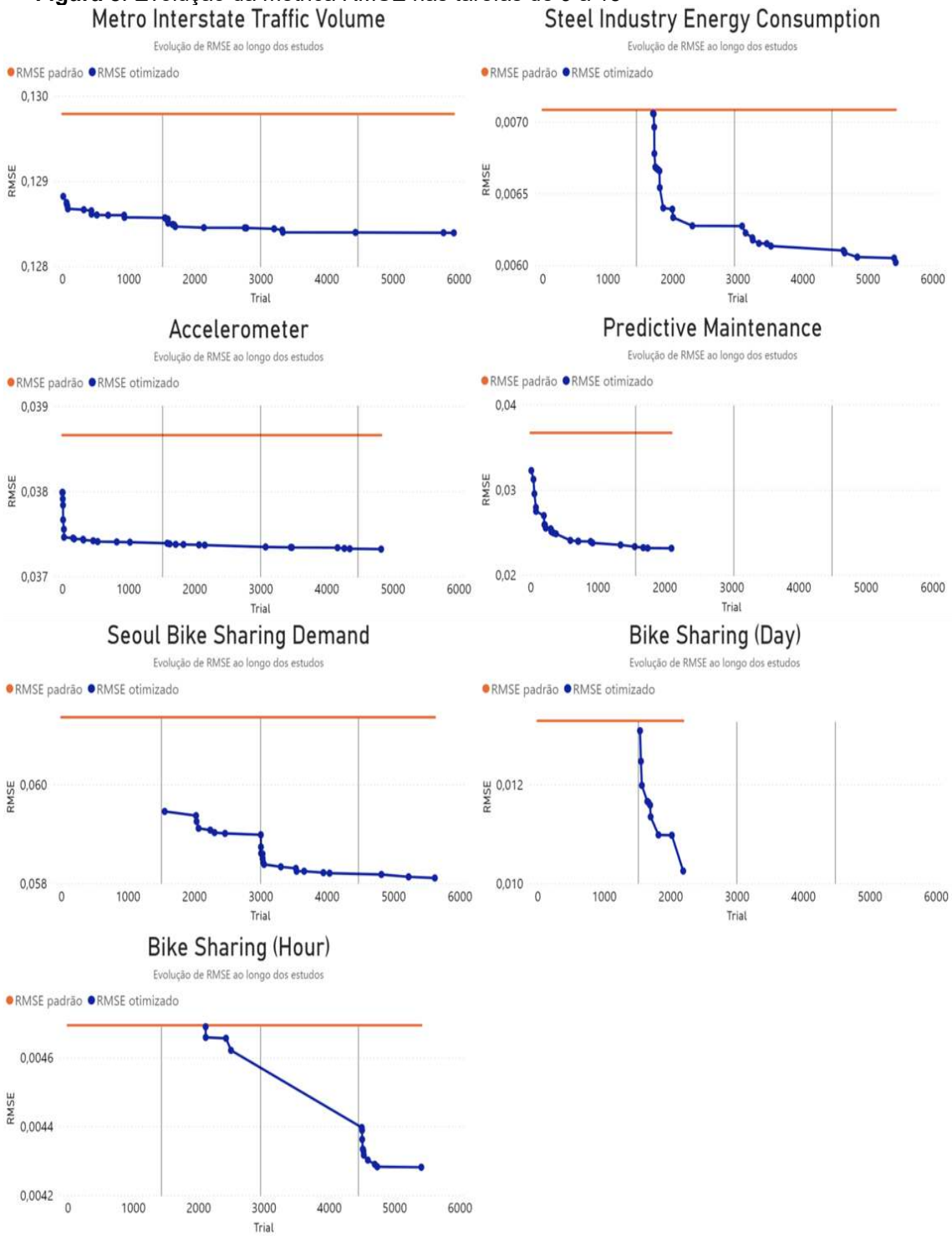
Figura 7. Evolução da métrica RMSE nas tarefas de 1 a 8



Fonte: elaboração própria (2026).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Figura 8. Evolução da métrica RMSE nas tarefas de 9 a 15



Fonte: elaboração própria (2026).

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

Ao final dos quatro ciclos, foram registrados valores de RMSE inferiores aos obtidos com as configurações padrão, conforme a Tabela 3.

Tabela 3. Comparação entre configurações padrão e melhores resultados encontrados

ID	*Dataset*	*RMSE com hiperparâmetros padrão*	*Melhor RMSE encontrado*	*Diminuição do RMSE*
01	Communities and Crime	0,1455097594	0,1294587404	11,03%
02	Productivity Prediction of Garment Employees	0,0910963492	0,0529790183	41,84%
03	Combined Cycle Power Plant	0,0412793049	0,0394047126	4,54%
04	Solar Flare (Common Flares)	0,1035761657	0,0833633365	19,51%
05	Solar Flare (Moderate Flares)	0,0840333003	0,0763343154	9,16%
06	Solar Flare (Severe Flares)	0,0483745421	0,0279074096	42,31%
07	Parkinsons Telemonitoring (Total)	0,1506086993	0,1390340153	7,69%
08	Parkinsons Telemonitoring (Motor)	0,1623243167	0,1474866716	9,14%
09	Metro Interstate Traffic Volume	0,1297878603	0,1283905287	1,08%
10	Steel Industry Energy Consumption	0,0070842752	0,0060175642	15,06%
11	Accelerometer	0,0386591962	0,0373198470	3,46%
12	Predictive Maintenance	0,0366795586	0,0231111574	36,99%
13	Seoul Bike Sharing Demand	0,0613538420	0,0581135042	5,28%
14	Bike Sharing (Day)	0,0132820243	0,0102541635	22,80%
15	Bike Sharing (Hour)	0,0046946870	0,0042814136	8,80%

Fonte: elaboração própria (2026).

As reduções registradas em relação às configurações padrão variaram entre 1,08% e 42,31%, com média de 15,91% após os quatro ciclos. A comparação entre o resultado final e o melhor resultado ao término das primeiras 1500 tentativas apresentou redução média de 4,24%, conforme a Tabela 4. Esses percentuais descrevem as comparações realizadas no estudo.

Tabela 4. Reduções registradas após o primeiro ciclo e após quatro ciclos

ID	*Dataset*	*Diminuição do RMSE com o primeiro estudo*	*Diminuição do RMSE com os 4 estudos*	*Diminuição do RMSE em relação ao fim do primeiro estudo*
01	Communities and Crime	9,27%	11,03%	1,94%
02	Productivity Prediction of Garment Employees	41,84%	41,84%	0,00%
03	Combined Cycle Power Plant	0,00%	4,54%	4,54%
04	Solar Flare (Common Flares)	19,47%	19,51%	0,05%
05	Solar Flare (Moderate Flares)	8,79%	9,16%	0,40%
06	Solar Flare (Severe Flares)	41,24%	42,31%	1,83%
07	Parkinsons Telemonitoring (Total)	7,00%	7,69%	0,74%
08	Parkinsons Telemonitoring (Motor)	9,10%	9,14%	0,04%
09	Metro Interstate Traffic Volume	0,94%	1,08%	0,14%

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

ID	Dataset	*Diminuição do RMSE com o primeiro estudo*	*Diminuição do RMSE com os 4 estudos*	*Diminuição do RMSE em relação ao fim do primeiro estudo*
10	Steel Industry Energy Consumption	0,00%	15,06%	15,06%
11	Accelerometer	3,25%	3,46%	0,22%
12	Predictive Maintenance	35,92%	36,99%	1,68%
13	Seoul Bike Sharing Demand	0,00%	5,28%	5,28%
14	Bike Sharing (Day)	0,00%	22,80%	22,80%
15	Bike Sharing (Hour)	0,00%	8,80%	8,80%

Fonte: elaboração própria (2026).

Os ganhos variaram entre as tarefas. Em *Productivity Prediction of Garment Employees*, o resultado não acrescentou redução ao melhor valor do primeiro ciclo. Já *Bike Sharing (Day)* e *Steel Industry Energy Consumption* apresentaram reduções posteriores ao primeiro ciclo. A mudança dos intervalos de busca é uma explicação possível para essa evolução, mas o delineamento não permite separar seu efeito daquele decorrente das tentativas adicionais.

4.3. Limitações e alcance dos resultados

A comparação entre um e quatro ciclos envolve orçamentos distintos de otimização. Não foi apresentada uma estratégia de controle com o mesmo número total de tentativas ou o mesmo tempo computacional. Assim, a redução adicional observada não demonstra, isoladamente, que a MOHOC seja mais eficiente do que continuar a busca original pelo mesmo orçamento.

A reutilização da partição de avaliação na função objetivo limita a interpretação dos valores como estimativas de generalização. Além disso, a seleção por correlação foi aplicada antes do particionamento, permitindo que informações do conjunto completo participassem da definição dos preditores. O ajuste dos normalizadores apenas no treino evita esse problema nessa etapa específica, mas não elimina a limitação da seleção anterior.

A utilização de partições aleatórias também deve ser considerada ao interpretar bases com dependências temporais ou observações agrupadas. As tarefas derivadas de uma mesma base compartilham características, e sua contagem não equivale à de amostras experimentais independentes. A média das reduções é uma síntese descritiva, sem teste de significância ou intervalo de confiança apresentado neste estudo.

As escolhas manuais dos intervalos, a ausência de mensuração do tempo de análise e a documentação incompleta das configurações por tarefa e ciclo restringem a reprodutibilidade e as conclusões sobre economia de recursos. Também não foram incorporadas variáveis exógenas além dos atributos das bases selecionadas. A contribuição deve ser entendida como uma demonstração aplicada de organização e interpretação dos registros de busca.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

5. CONSIDERAÇÕES FINAIS

Foi apresentado e aplicado um procedimento de refinamento manual do espaço de busca do Optuna para o XGBoost em quinze tarefas de regressão derivadas de onze bases públicas. A ferramenta MOHOC organiza registros de tentativas em visualizações que apoiam a interpretação dos resultados e a definição de intervalos para ciclos posteriores.

Ao final dos quatro ciclos, os valores registrados apresentaram redução média de 15,91% em relação às configurações padrão e de 4,24% em relação ao primeiro ciclo. Esses resultados indicam melhora da métrica acompanhada durante a busca nas condições estudadas. Não permitem afirmar superioridade geral do procedimento, obtenção de um ótimo global ou desempenho equivalente em dados externos.

A utilização prática do procedimento pressupõe conhecimento do modelo, dos limites válidos dos hiperparâmetros e das características dos dados. Como continuidade da pesquisa, recomenda-se comparar estratégias com orçamentos equivalentes, separar uma avaliação final independente e documentar regras de atualização dos intervalos e o tempo de intervenção humana. Essas recomendações delimitam estudos futuros e não integram os experimentos aqui apresentados.

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL

Foi utilizada assistência de inteligência artificial na revisão textual, tradução, organização e formatação do manuscrito. Os dados experimentais e os resultados numéricos não foram gerados por inteligência artificial. A responsabilidade pelo conteúdo científico permanece com o autor.

REFERÊNCIAS

AKIBA, T.; SANO, S.; YANASE, T.; OHTA, T.; KOYAMA, M. Optuna: a next-generation hyperparameter optimization framework. In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 25., 2019. **Proceedings [...]**. ACM, 2019. p. 2623–2631. DOI: https://doi.org/10.1145/3292500.3330701.

ALQUTHAMI, T.; ZULFIQAR, M.; KAMRAN, M.; MILYANI, A. H.; RASHEED, M. B. A performance comparison of machine learning algorithms for load forecasting in smart grid. **IEEE Access**, v. 10, p. 48419–48433, 2022. DOI: https://doi.org/10.1109/access.2022.3171270.

AL-RAKHAMI, M.; GUMAEI, A.; ALSANAD, A.; ALAMRI, A.; HASSAN, M. M. An ensemble learning approach for accurate energy load prediction in residential buildings. **IEEE Access**, v. 7, p. 48328–48338, 2019. DOI: https://doi.org/10.1109/access.2019.2909470.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.

ARDEN, F.; SAFITRI, C. Hyperparameter tuning algorithm comparison with machine learning algorithms. In: INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY, INFORMATION SYSTEMS AND ELECTRICAL ENGINEERING (ICITISEE), 6., 2022. **Proceedings [...]**. IEEE, 2022. p. 183–188. DOI: https://doi.org/10.1109/icitisee57756.2022.10057630.

BOEHMKE, B. C. **Data Wrangling with R**. Cham: Springer, 2016. DOI: https://doi.org/10.1007/978-3-319-45599-0.

CHEN, T.; GUESTRIN, C. XGBoost: a scalable tree boosting system. In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 22., 2016. **Proceedings [...]**. ACM, 2016. p. 785–794. DOI: https://doi.org/10.1145/2939672.2939785.

DEMIDOVA, L. A.; FILATOV, A. V. Optimization of hyperparameters with constraints on time and memory for the classification model of the hard drives states. In: INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGIES (InfoTech), 2022. **Proceedings [...]**. IEEE, 2022. p. 1–4. DOI: https://doi.org/10.1109/infotech55606.2022.9897074.

PRASAD, S. S.; DEO, R. C.; DOWNS, N.; IGOE, D.; PARISI, A. V.; SOAR, J. Cloud affected solar UV prediction with three-phase wavelet hybrid convolutional long short-term memory network multi-step forecast system. **IEEE Access**, v. 10, p. 24704–24720, 2022. DOI: https://doi.org/10.1109/ACCESS.2022.3153475.

SEPPE, F. R. **MOHOC: Manual Optuna Hyperparameter Optimization Cycle**. Versão 1.0.0. Zenodo, 2026. Software. DOI: https://doi.org/10.5281/zenodo.24096106.

TAN, M.; YUAN, S.; LI, S.; SU, Y.; LI, H.; HE, F. Ultra-short-term industrial power demand forecasting using LSTM based hybrid ensemble learning. **IEEE Transactions on Power Systems**, v. 35, n. 4, p. 2937–2948, 2020. DOI: https://doi.org/10.1109/tpwrs.2019.2963109.

YANG, L.; SHAMI, A. On hyperparameter optimization of machine learning algorithms: theory and practice. **Neurocomputing**, v. 415, p. 295–316, 2020. DOI: https://doi.org/10.1016/j.neucom.2020.07.061.

ZHANG, D.; GONG, Y. The comparison of LightGBM and XGBoost coupling factor analysis and prediagnosis of acute liver failure. **IEEE Access**, v. 8, p. 220990–221003, 2020. DOI: https://doi.org/10.1109/access.2020.3042848.

Este artigo é publicado em acesso aberto (Open Access) sob a licença Creative Commons Atribuição 4.0 Internacional (CC-BY), que permite uso, distribuição e reprodução irrestritos em qualquer meio, desde que o autor original e a fonte sejam creditados.