

INVESTIGAÇÃO DA MELHORIA NA FUNÇÃO OBJETIVO DA OTIMIZAÇÃO BAYESIANA

OCTÁVIO RODRIGUES FILHO¹; CELSO RICARDO CALDEIRA RÉGO²;
MAURÍCIO JEOMAR PIOTROWSKI³

¹*Universidade Federal de Pelotas – octavio.fisica.lab@gmail.com*

² *Karlsruhe Institute of Technology – celso.rego@kit.edu*

³*Universidade Federal de Pelotas – mauriciomjp@gmail.com*

1. INTRODUÇÃO

Os processos físicos e químicos são frequentemente descritos por modelos matemáticos, que desempenham um papel crucial no poder preditivo em ambas as áreas. Em alguns casos, como na Teoria do Funcional da Densidade (KOHN, 1964), busca-se minimizar a energia com base na densidade eletrônica. Em outros contextos, a Otimização Bayesiana (BO) (FRAZIER, 2018) tem sido utilizada para a calibração e seleção de funcionais híbridos (VASCONCELOS et al., 2020). Um exemplo prático é a exploração acelerada de catalisadores heterogêneos para a hidrogenação de CO₂ (RAMIREZ et al., 2024), onde os objetivos são a maximização da conversão de CO₂ e da seletividade de metanol, ao mesmo tempo que se minimiza o custo dos metais e a seletividade de metano. Nesses casos, funções matemáticas são empregadas para descrever o sistema, com o objetivo de minimizar, maximizar, calibrar ou prever esses comportamentos.

Quando nos deparamos com a ausência de uma função ou modelo claro que governe o sistema, enfrentamos o que se chama de "problemas de caixa preta" (PHAN-TRONG et al., 2023). Para abordar esses problemas, pode-se recorrer a regressores e classificadores (LOH, 2011), que são modelos matemáticos usados para prever valores com base em entradas, ajustando-se de acordo com o erro de previsão. Um exemplo popular é o *Extreme Gradient Boosting* (XGB) (CHEN et al., 2016), um regressor da família *Gradient Boosting*, que se diferencia das *Random Forests* (BREIMAN, 2001) ao criar árvores de decisão sequenciais, nas quais cada nova árvore corrige o erro das anteriores. Embora o XGB permita a previsão eficiente de conjuntos de dados, ele pode ser insuficiente para definir o valor ótimo dentro do espaço amostral e não impõe restrições ao mesmo, o que pode levar ao *overfitting* (YANG, 2019), resultando em uma validação menos eficaz.

Para identificar os melhores valores dentro de um espaço amostral e restringir esse espaço, a BO oferece uma solução. A BO é um modelo de aprendizado de máquina baseado em amostragem estatística, que busca maximizar ou minimizar uma função objetivo (FRAZIER, 2018). Na BO, existe o modelo substituto ($f(x)$), que é nossa hipótese de modelo matemático, e a função de aquisição ($g(x)$), que determina, a partir de $f(x)$ e da função objetivo, o melhor valor dentro do espaço amostral. Normalmente, utiliza-se um Processo Gaussiano (GP) (WILLIAMS et al., 1995) como regressor.

O objetivo deste trabalho é apresentar uma metodologia alternativa, na qual o XGB é utilizado como função objetivo para criar a relação entre entrada e saída dos dados, enquanto o GP é empregado para modelar as incertezas nas previsões do XGB. O XGB é treinado uma única vez e não é re-treinado, servindo como uma função fixa dentro da BO, o que resulta em uma menor extrapolação e evita a superotimização. Além disso, empregamos um banco de dados disponível na literatura (RAMIREZ et al., 2024) para validar essa abordagem e investigar possíveis ocorrências de superotimização.

2. METODOLOGIA

Utilizando a linguagem de programação Python, desenvolvemos um *script* capaz de processar os dados e realizar a previsão de entrada e saída utilizando o modelo XGB. Após essa etapa, o módulo de previsão gerado pelo XGB é entregue à função objetivo. A BO atua na identificação dos melhores valores dentro de um espaço amostral restrito, maximizando ou minimizando a função objetivo. Nesse contexto, estamos lidando com uma otimização multiobjetivo, onde os parâmetros da função objetivo e sua previsão são os alvos, ao invés de focar apenas na saída final da função.

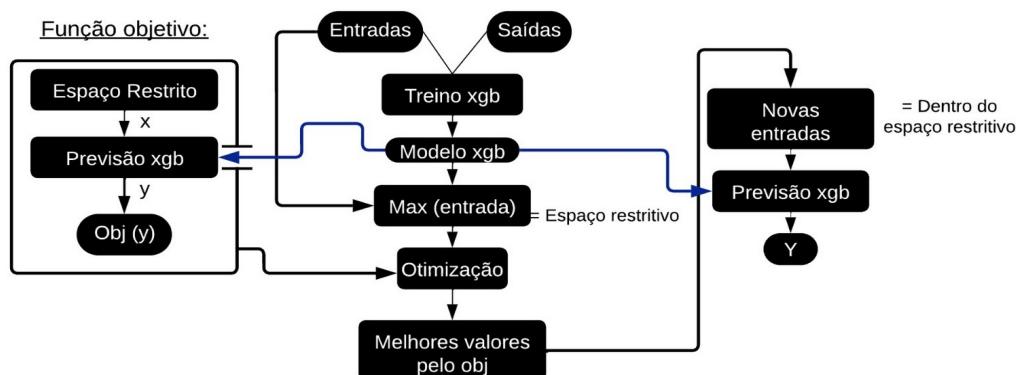


Figura 1: Fluxograma representativo das operações do código, após o treino é efetuado a previsão dentro da função objetivo e coletado os melhores candidatos que minimizam e/ou maximizam a saída da função objetivo.

A função objetivo é formulada com um espaço restrito, orientado pelo valor máximo que as entradas podem alcançar, como ilustrado na Figura 1. Esse espaço recebe o total de parâmetros relacionados às entradas, sendo a saída o valor ajustado pela otimização. Dentro desse processo, os parâmetros estabelecem a relação entre entrada e saída por meio da previsão gerada pelo XGB. A equação da função objetivo é dada por:

$$OBJ(Parâmetros) = - \sum (y_i) \quad (1)$$

Para os valores de previsão que minimizam ou maximizam a saída da equação (1), obtemos os resultados correspondentes à otimização. No caso da minimização, o valor deve ser negativo, enquanto na maximização, o valor é indicado com sinal positivo. Como estamos lidando com um problema multiobjetivo em y_i , é necessário especificar quais parâmetros são minimizados e quais são maximizados. Para a Otimização Bayesiana, escolhemos a função de aquisição "Probabilidade de Melhoria" (SNOEK et al., 2012), que se mostrou adequada para nosso propósito.

Para validar a metodologia de implementação do XGB na equação (1), decidimos maximizar a conversão de CO₂ e a seletividade de metanol, ao mesmo tempo em que minimizamos o custo dos metais e a seletividade de metano, em consonância com a literatura (RAMIREZ et al., 2024). A ideia principal desta metodologia é replicar o processo descrito na literatura, mas utilizando uma abordagem diferente, onde não possuímos as equações governantes do sistema e forçamos a criação da relação de entrada e saída através do XGB, seguido da otimização.

3. RESULTADOS E DISCUSSÃO

O processo de otimização foi conduzido ao longo de quatro gerações, com o índice inicial começando na segunda geração. A primeira geração foi utilizada para estabelecer os limiares iniciais no script. O modelo XGB apresentou um erro quadrático médio de 0,0117, o que nos permitiu selecionar os melhores resultados com base nos menores erros, conforme diferentes conjuntos de parâmetros.

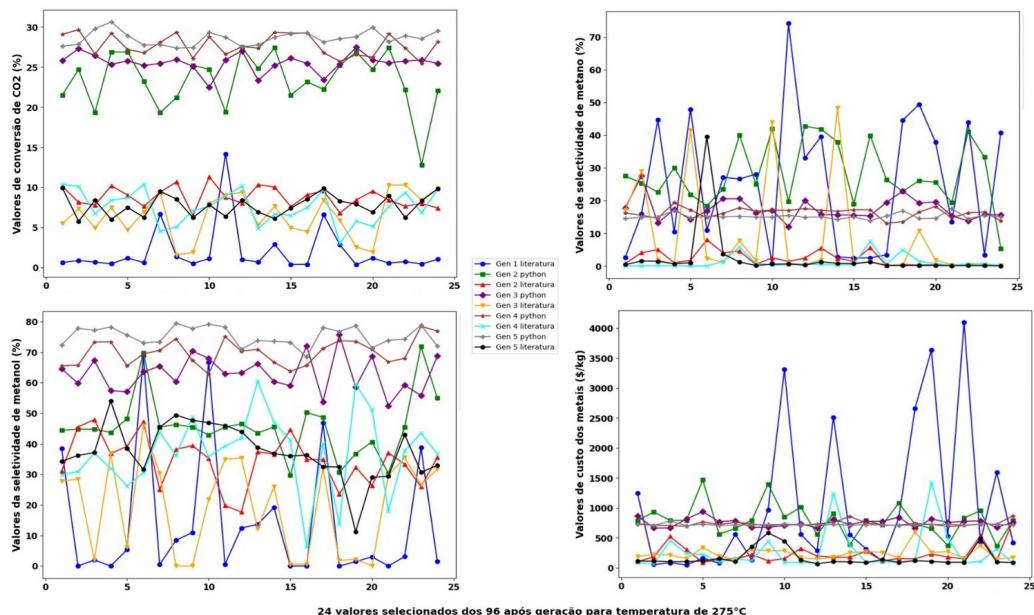


Figura 2: Comparação dos resultados da otimização com dados da literatura. O eixo vertical representa os valores obtidos para as métricas de conversão de CO₂ e seletividade de metanol, enquanto o eixo horizontal apresenta os 24 melhores valores obtidos para a temperatura de 275°C.

A análise dos resultados apresentados na Figura 2 indicou uma melhoria significativa na conversão de CO₂. Comparado com os valores da literatura, que reportam uma conversão de aproximadamente 14%, conseguimos alcançar até 30% em alguns casos. Esse aumento demonstra a eficácia da metodologia proposta. Além disso, a seletividade de metanol foi otimizada eficientemente em três gerações consecutivas, confirmando a capacidade do modelo em ajustar esses parâmetros críticos.

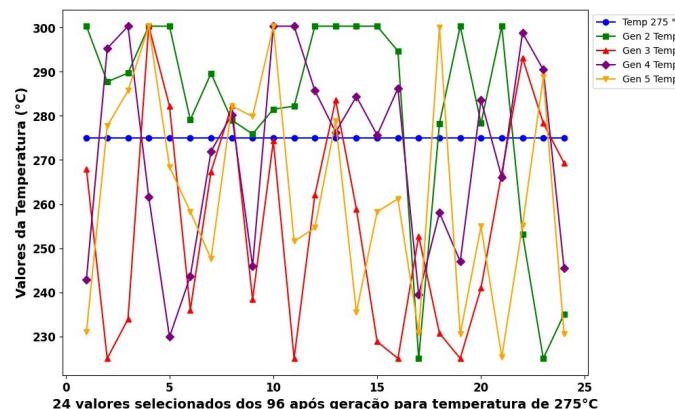


Figura 3: Relação da temperatura original e temperaturas após otimização.

Em contrapartida, os parâmetros relacionados às minimizações apresentaram comportamento mais estável, com preferência por valores constantes. Isso sugere que, enquanto as maximizações responderam bem à otimização, os critérios de minimização foram mais conservadores, mantendo-se em um patamar controlado.

A Figura 3 destaca os critérios de temperatura que foram ajustados ao longo das gerações, o que contribuiu para a melhoria dos resultados observados. Esses ajustes nos permitiram reproduzir novas gerações com base nas modificações térmicas aplicadas, favorecendo a otimização dos parâmetros alvo.

4. CONCLUSÕES

Concluímos que a modificação da função objetivo implementada com o modelo XGB apresentou uma melhoria significativa, fornecendo uma solução eficaz para problemas de caixa preta, nos quais a matemática subjacente ao conjunto de dados é desconhecida propositalmente pra simular o problema. Além de prever os resultados, o método também permite a otimização dos mesmos. Os valores obtidos se diferenciam da literatura devido à escolha do "melhor valor entre os melhores valores", o que garante uma superotimização, mas pode comprometer a qualidade dos dados. Em termos gerais, alcançamos valores elevados para as maximizações e valores constantes para as minimizações, constatando que a abordagem é viável para a exploração de problemas de caixa preta.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- KOHN, Walter. Inhomogeneous electron gas. *Phys. Rev.*, v.136, p.B864, 1964.
- FRAZIER, Peter I. A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811*, p.1, 2018.
- VASCONCELOS, Hugo; ALMEIDA, Rodrigo. Bayesian optimization for calibrating and selecting hybrid-density functional models. *J. Phy. Chem. A*, v.124, p.4053, 2020.
- RAMIREZ, Adrian; LAM, Erwin. Accelerated exploration of heterogeneous CO₂ hydrogenation catalysts by bayesian-optimized high-throughput and automated experimentation. *Chem Catal.*, v.4, p.1, 2024.
- LOH, Wei-Yin. Classification and regression trees. *Wiley Interd. Rev.: Data Mining and Knowledge Discovery*, v.1, p.14, 2011.
- YANG, Xue. An overview of overfitting and its solutions. *J. Phys.: Conf. Series*, v.1168, p.022022, 2019.
- WILLIAMS, Christopher K. I.; RASMUSSEN, Carl E. Gaussian processes for regression. *Adv. Neural Inf. Proc. Syst.*, v.8, 1995.
- CHEN, Tianqi; GUESTRIN, Carlos. Xgboost: A scalable tree boosting system. In: *Proc. 22nd ACM SIGKDD Int. Conf. Know. Disc. Data Min.* New York, NY, USA: ACM, 2016. p.785-794.
- BREIMAN, Leo. Random forests. *Machine Learning*, v.45, p.5, 2001.
- SNOEK, Jasper; LAROCHELLE, Hugo; ADAMS, Ryan P. Practical bayesian optimization of machine learning algorithms. *Adv. Neural Inf. Proc. Syst.*, v.25, 2012.
- PHAN-TRONG, Dat; GROSS, Samuel; TRAN THANH, Long. Neuralbo: A blackbox optimization algorithm using deep neural networks. *Neurocomput.*, v.559, p.126776, 2023.