MODELAGEM MOLECULAR E IA PARA A PREDIÇÃO DE PROPRIEDADES FÍSICO-QUÍMICAS DE COMBUSTÍVEIS

OLIVEIRA, Eric Soares[1]; PIRES, Vinicius Santana[1]; SANCHES-NETO, Flávio O.[1]*
[1]Instituto Federal de Goiás, IFG-Câmpus Valparaíso de Goiás *flavio.neto@ifg.edu.br

Este estudo explora detalhadamente a aplicação de métodos avançados de aprendizado de máquina combinados com relações quantitativas entre estrutura e propriedade (ML-QSPR) para a predição precisa de propriedades físico-químicas de diversos tipos de combustíveis. A motivação principal deste trabalho é aprimorar a capacidade preditiva em relação a propriedades críticas como o número de cetano e a temperatura de fusão, que são fundamentais para o desempenho e a eficiência energética dos combustíveis. Inspirado por procedimentos metodológicos adotados em estudos recentes na área, o presente trabalho adaptou e refinou essas técnicas para um conjunto de dados específico, composto por uma variedade de compostos combustíveis com diferentes estruturas moleculares. O processo iniciou-se com uma etapa minuciosa de coleta de dados, seguida por um rigoroso processo de limpeza e tratamento, visando eliminar inconsistências, dados ausentes e outliers que pudessem comprometer a qualidade dos modelos preditivos. Diversos algoritmos de aprendizado de máquina foram empregados para modelagem e análise dos dados, incluindo métodos tradicionais como a regressão linear e técnicas mais complexas como Random Forest Regression e XGBoost. Cada um desses algoritmos foi cuidadosamente configurado e otimizado por meio de técnicas de validação cruzada e ajuste de hiperparâmetros, com o objetivo de maximizar a precisão das previsões. As métricas de avaliação utilizadas para comparar o desempenho dos modelos incluíram o erro quadrático médio (RMSE), o coeficiente de determinação (R^2) e a precisão percentual média. Os resultados obtidos demonstraram que o modelo Random Forest Regression apresentou o melhor desempenho geral, superando os demais algoritmos em termos de precisão e robustez das previsões. No entanto, observou-se que as precisões alcançadas ainda ficaram aquém das reportadas na literatura especializada, indicando a necessidade de melhorias adicionais. Este trabalho contribui significativamente para a compreensão das capacidades e limitações de diferentes algoritmos de aprendizado de máquina na predição de propriedades críticas de combustíveis. As análises realizadas fornecem insights valiosos para futuras pesquisas, sugerindo que a combinação de métodos de aprendizado profundo com técnicas de engenharia de características pode potencialmente melhorar a precisão dos modelos. Além disso, as conclusões deste estudo têm implicações práticas importantes, fornecendo uma base sólida para otimizações futuras e aplicações em química computacional e ambiental, especialmente no desenvolvimento de combustíveis mais eficientes e sustentáveis.

Palavras-chave: ML-QSPR, Propriedades Físico-Químicas, Combustíveis, Aprendizado de Máquina, Modelagem Molecular.

Agradecimentos: O presente trabalho foi realizado com apoio do Instituto Federal de Goiás (n° 19/2024). Flávio Olimpio Sanches Neto agradece aos auxílios recebidos pela FAPEG, CAPES e CNPq.