cbecimat26
26° Congresso Brasileiro de Engenharia e Ciência dos Materiais
22 a 26 de novembro de 2026
Costão do Santinho
Florianópolis-SC

Comparison of machine learning methods prediction for TiO_2 properties

Natan Bueno Ungethuem, Júllia de Marins Sbardelotto, Vinícius Mayer Fontana, Colombo Freisleben, Vânia Caldas de Sousa
Federal University of Rio Grande do Sul – UFRGS
Programa de pós-graduação em engenharia de minas, metalúrgica e materiais – PPGE3M

Introduction

This work employs Machine Learning (ML) methods, such as: neural networks, Random Forest, XGBoost; and generic Large Language Models (LLM) to predict properties of TiO_2, a oxide ceramic material, applied to sunscreen.

Goals

1. Demonstrate that AI agent orchestration, ML and LLM are powerful tools to accelerate the design of new materials
2. Verify if its possible to use ML and LLM to predict cell viability of TiO_2 used to sunscreen

Methodology

The methodology followed a three-step process: data collection, training, and prediction. Data collection use three different AI agents. Models to trainings were ML and generic LLM models.

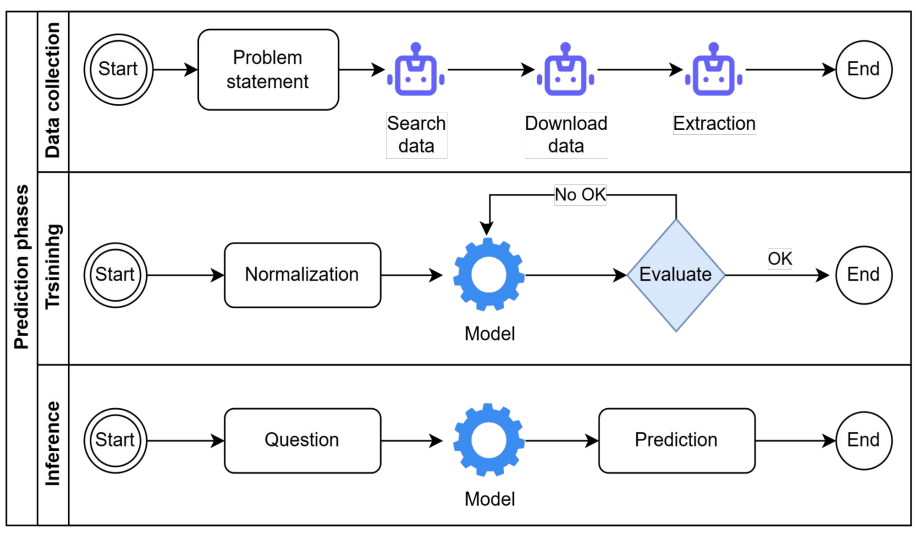


Figure 1

Results

The LLM tools proved highly valuable to search data in scientific articles, delivering results similar to human search in a fraction of the time: 6 months reduced to just 3 days.

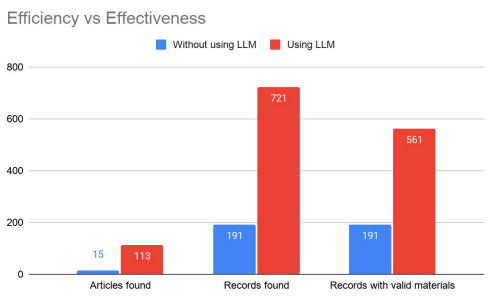


Figure 2

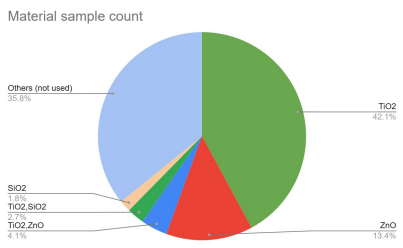


Figure 3

We tried to create a Neural Network model, but it was discarded due to a degenerate behavior.

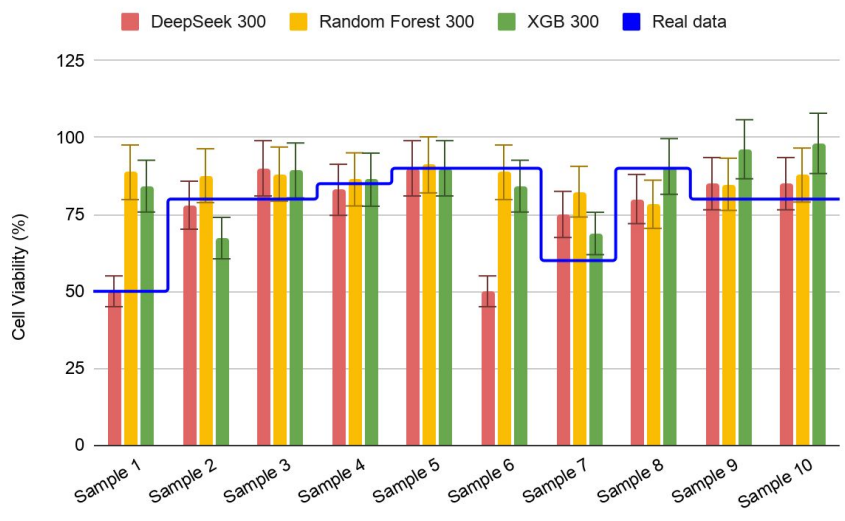


Figure 4

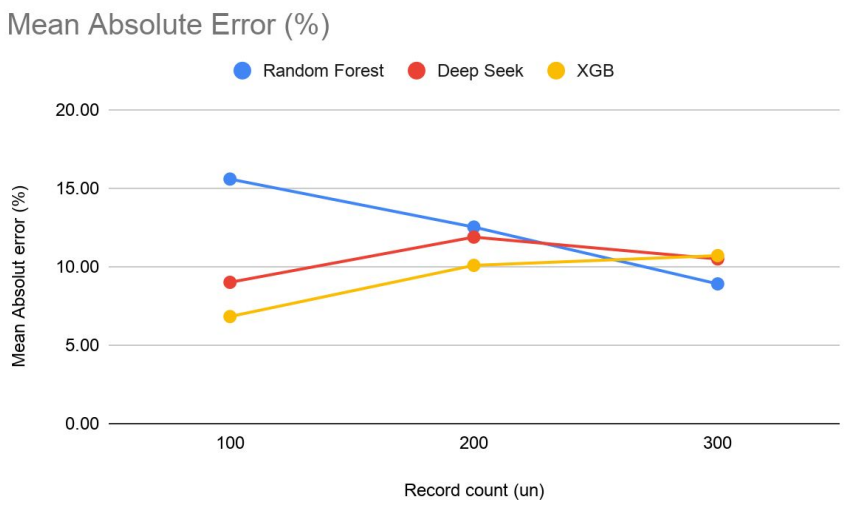


Figure 5

Conclusion

The DeepSeek and XGBoost models demonstrated better accuracy in scenarios with few records. With an increased volume of data, which introduced greater randomness, an improvement was observed in the general trend of the predictions. With large volumes of data, the accuracy of the models begins to converge, making DeepSeek a more advantageous option, as it requires less development time for creating the simulator.
Finally, the use of orchestrated AI agents proved to be a substantial benefit in optimizing the search for records to be tested, reducing the data preparation time.

References and acknowledgement

NASER, M. Z. **A review of machine learning with small and limited data**. Journal of Big Data, v. 13, n. 18, 2026. DOI: https://doi.org/10.1186/s40537-025-01346-9.
ZHANG, Yangfan et al. **Machine Learning for Designing Perovskites and Perovskite-Inspired Solar Materials: Emerging Opportunities and Challenges**. Advanced Science, v. 13, e74952, 2026. DOI: https://doi.org/10.1002/advs.74952. Disponível em: https://advanced.onlinelibrary.wiley.com/doi/pdf/10.1002/advs.74952
CHEN, Shengbin; ZHANG, Wenming; LUO, Rui; ZHAO, Yidong; YANG, Yang; ZHANG, Bing; LU, Qiang; HU, Bin. **System energy and band gap prediction of titanium dioxide based on machine learning. Journal of Molecular Structure**, v. 1307, n. 137934, 2024. DOI: 10.1016/j.molstruc.2024.137934. Disponível em: https://www.sciencedirect.com/science/article/pii/S0022286024004575

CAPES


Natan Bueno Ungethuem


Júllia de Marins Sbardelotto


Vinícius Mayer Fontana


Lucas Colombo Freisleben


Vânia Caldas de Sousa