

## Métodos de minería de datos y aprendizaje automático para optimizar los fotocatalizadores para la producción de hidrógeno

Baum, Fábio; Pereira, Thiago Rafael

Universidade Federal da Integração Latino-Americana, Foz do Iguaçu, Brasil

fabio.baum@unila.edu.br

Área temática: G. Aplicaciones de nanomateriales en ambiente, energía, agro, alimentos y catálisis

Según el informe de 2026 del Proyecto Global del Carbono, se proyecta que las emisiones globales de CO<sub>2</sub> de origen fósil alcancen un récord de 38.100 millones de toneladas, un aumento del 1,1 % con respecto a 2024 [1]. Esta alarmante tendencia subraya la urgente necesidad de alternativas energéticas sostenibles. La producción de hidrógeno mediante la disociación o reformado del agua por fotocatalisis se presenta como una solución prometedora para una matriz energética limpia. Sin embargo, estos sistemas aún carecen de la eficiencia necesaria para la producción de hidrógeno a gran escala. Para superar estas limitaciones, los enfoques de aprendizaje automático (ML) ofrecen la oportunidad de aprovechar los datos de la literatura para la optimización de fotocatalizadores para la producción de hidrógeno mediante luz solar.

En este estudio, se construyó un conjunto de datos exhaustivo extrayendo información de la literatura sobre fotocatalisis mediante la API DeepSeek para automatizar la recuperación de información. El conjunto de datos obtenido contiene alrededor de 2600 observaciones, con 50 características extraídas de 15 000 artículos de la literatura. Los siguientes modelos de regresión basados en árboles se entrenaron utilizando este conjunto de datos: árboles de decisión, bosques aleatorios, potenciación de gradiente, XGBoost y árboles de regresión aditiva bayesiana (BART). El rendimiento del modelo se evaluó mediante el error absoluto medio (MAE), el error cuadrático medio (RMSE) y el coeficiente de determinación (R<sup>2</sup>). Al identificar descriptores clave de fotocatalizadores y parámetros de reacción, esta metodología minimiza la dependencia de experimentos de ensayo y error costosos y que consumen mucho tiempo. Por lo tanto, este trabajo contribuye a la transformación digital del sector de las energías renovables, ofreciendo una estrategia más rápida y rentable para optimizar la producción de hidrógeno mediante energía solar.

### REFERENCIAS

1. Friedlingstein, P.; Jones, M.W.; O'Sullivan, M. et al., *Earth System Science Data* 17 (2025) 965-1039