

DESARROLLO DE UN MODELO DE APRENDIZAJE PROFUNDO PARA LA PREDICCIÓN DEL ANCHO DE BANDA EN MATERIALES TERNARIOS BASADOS EN ZNO

M EN C. GABRIEL GARCIA ZAMBRANO, DR. ARTURO ISAÍAS MARTÍNEZ ENRÍQUEZ, DR. DANIEL OLGUÍN MELO RITO, DR. FELIPE MONDACA ESPINOZA.

RESUMEN

Determinar el bandgap de Óxidos Metálicos Semiconductores es esencial para diseñar dispositivos optoelectrónicos, pero los métodos de síntesis y caracterización resultan costosos y requieren tiempo de optimización. Para ofrecer una alternativa eficiente previa, desarrollamos un modelo de aprendizaje profundo que predice el bandgap en materiales ternarios basados en ZnO.

Compilamos una base de datos de cientos de archivos cristalográficos (.cif) de diversas fuentes. Usando Python, extrajimos propiedades relevantes y consolidamos los datos en un archivo de valores separado por comas (.csv) enriquecido con información de la literatura y cálculos DFT con el funcional HSE06.

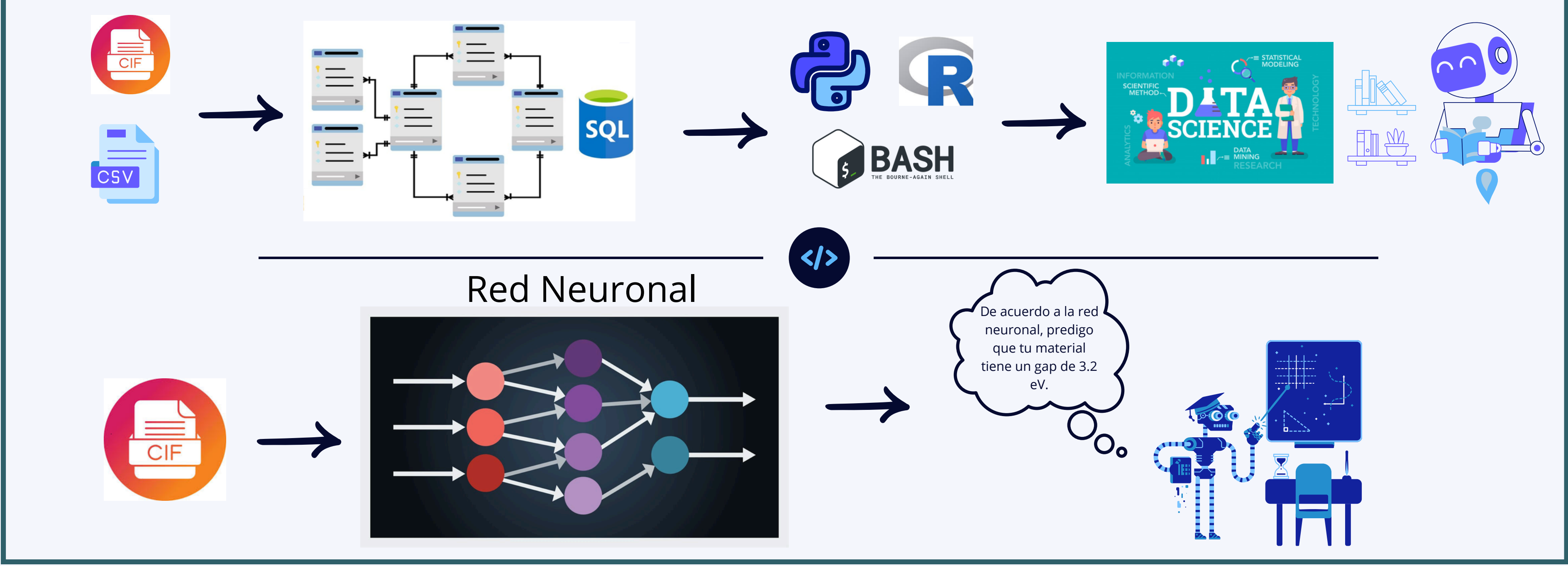
Una vez limpia y consolidada la información se desarrollo una base de datos estructurada (MySQL) la cuál incluye propiedades físicas, químicas y electrónicas. Creamos una **variable Objetivo** denominada "GAP", la cuál incluye el bandgap óptico reportada en la literatura.

Desarrollamos un modelo de red neuronal utilizando herramientas de scikit-learn, pandas y Keras. Seleccionamos las diez variables más influyentes tras analizar correlaciones. Optimizamos los hiperparámetros mediante validación cruzada y búsqueda en malla, mejorando el rendimiento. El modelo mostró mejor R^2 (0.87) y menor RMSE (0.92) que modelos basados en XGBoost (Machine Learning) demostrando la eficiencia de las redes neuronales artificiales.

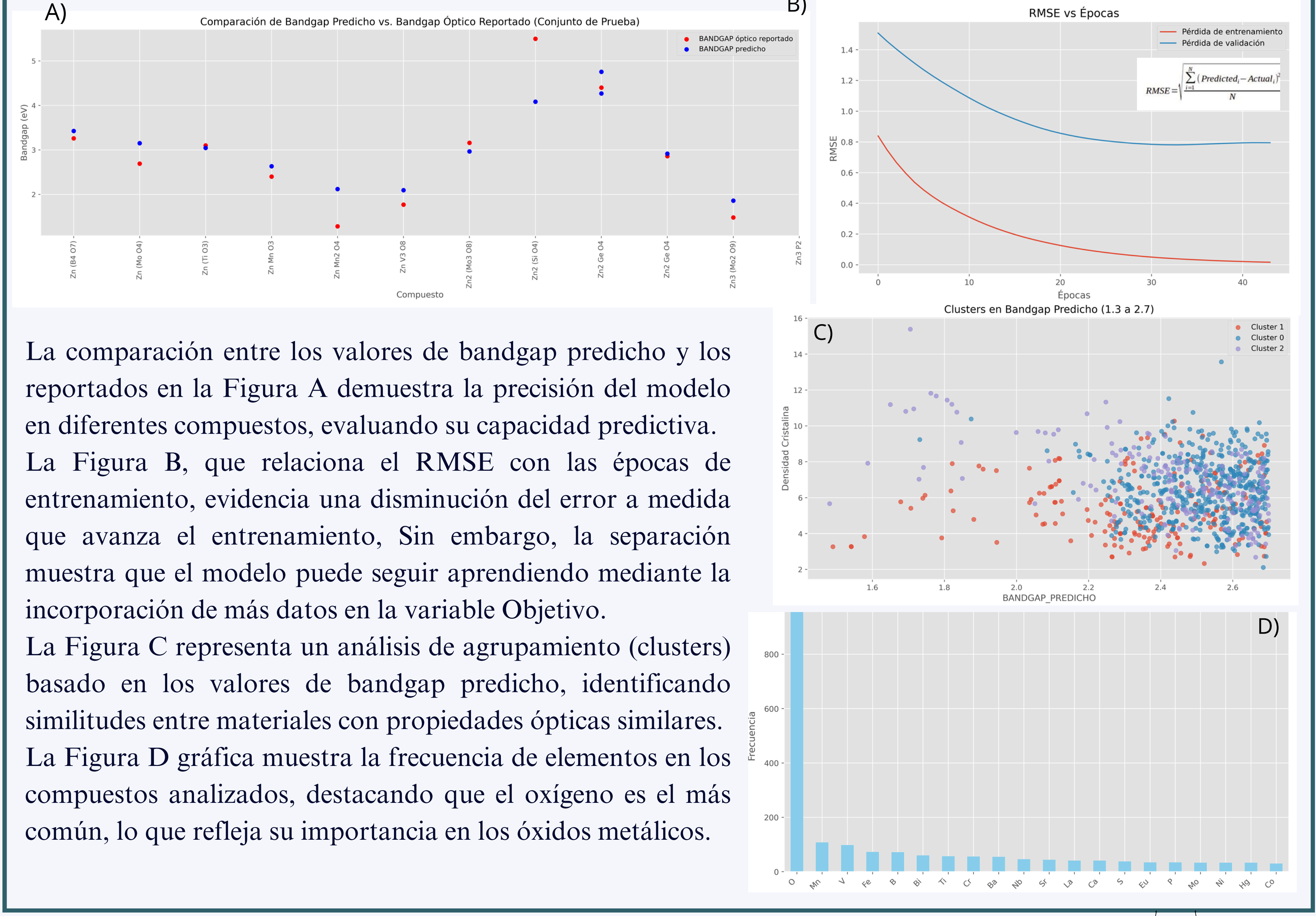
Este trabajo demuestra la eficacia de predecir el bandgap de materiales ternarios basados en ZnO mediante aprendizaje profundo, reduciendo costos y tiempos. La metodología puede adaptarse para predecir otras propiedades, contribuyendo al desarrollo de nuevos materiales y promoviendo la integración de la inteligencia artificial en el diseño de Materiales.

Palabras Clave: Bandgap, Aprendizaje profundo, Redes neuronales, Predicción de propiedades de materiales, Óxido de Zinc (ZnO), Teoría de los Funcionales de la Densidad (DFT)

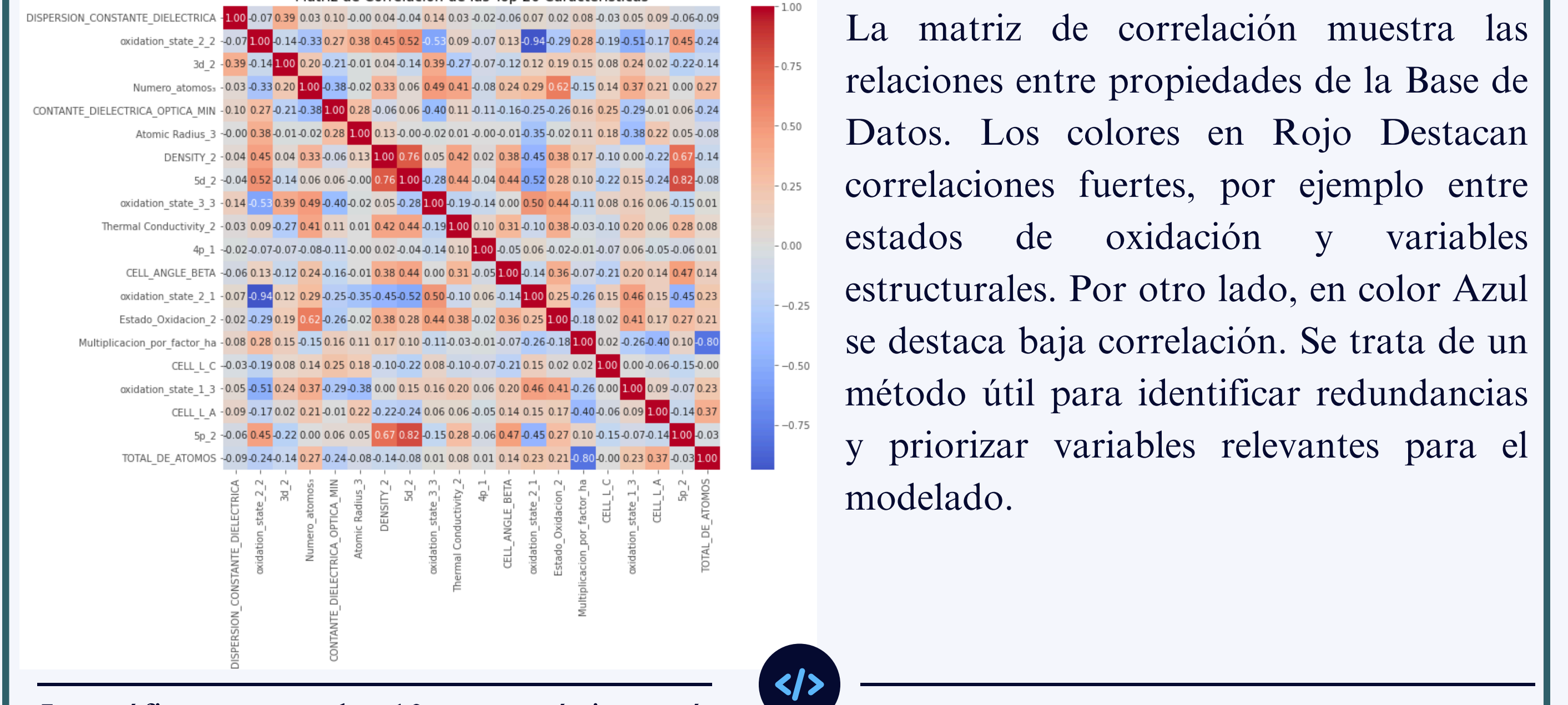
METODOLOGÍA



DISCUSIÓN



RESULTADOS



La gráfica muestra las 10 características más relevantes en un modelo predictivo, clasificadas por su importancia medida mediante el puntaje F Score. La constante dieléctrica óptica mínima es la más significativa, seguida de la constante dieléctrica efectiva y características electrónicas como 3d₂ y estados de oxidación. Estas variables son cruciales para explicar el comportamiento del modelo, así como para seleccionar las neuronas de entrada.

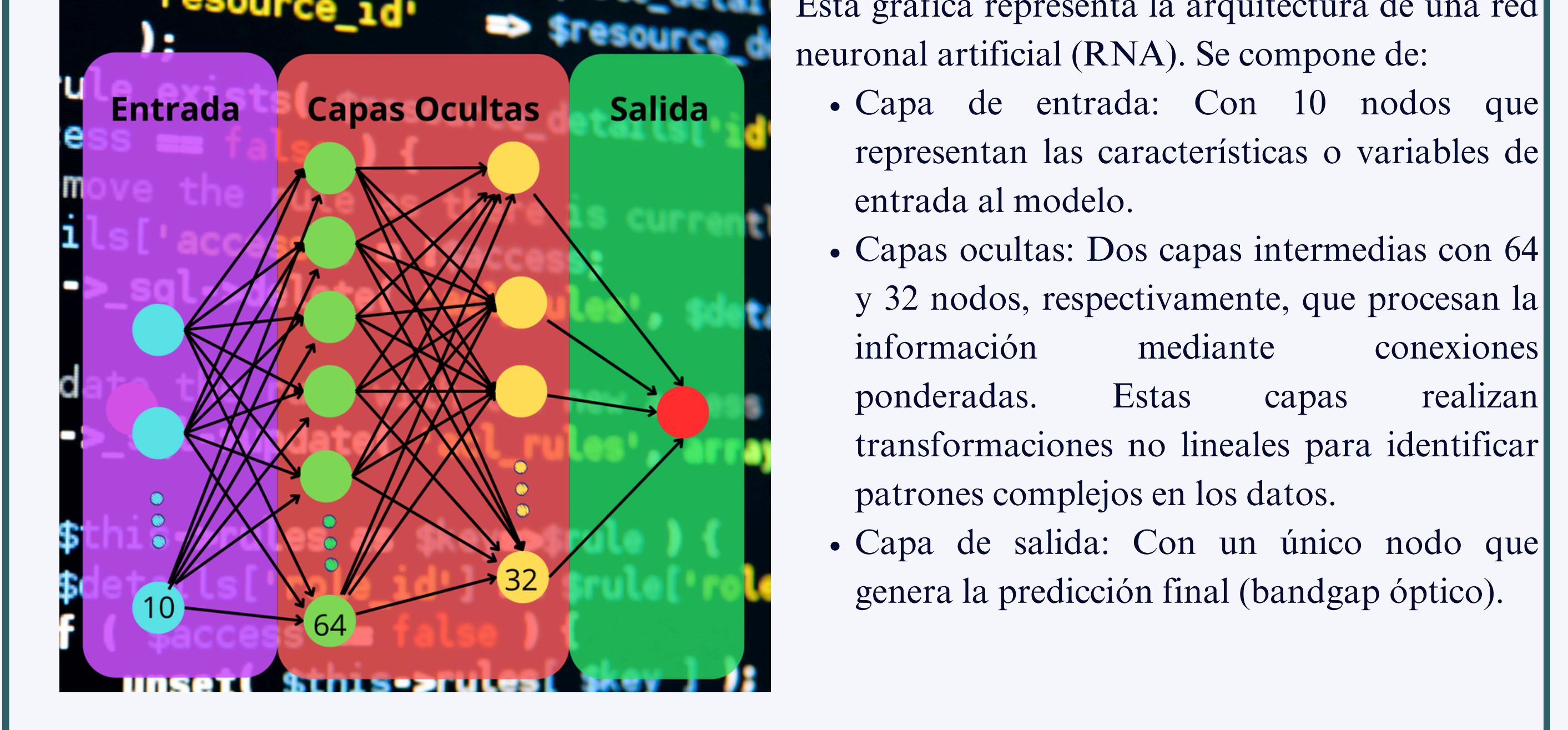
CONCLUSIÓN

En este estudio se demuestra la eficacia de predecir el bandgap óptico en materiales ternarios basados en ZnO mediante un modelo de aprendizaje profundo, optimizando así la investigación de materiales para aplicaciones optoelectrónicas. La combinación de datos extraídos de archivos cristalográficos y literatura, cálculos DFT con el funcional HSE06 y una base de datos estructurada permitió desarrollar un modelo de red neuronal con alto desempeño ($R^2 = 0.87$, $RMSE = 0.92$). Este modelo superó a enfoques basados en algoritmos tradicionales como XGBoost, destacando la capacidad de las redes neuronales para identificar patrones complejos en datos multidimensionales.

La metodología incluyó la selección de las variables más influyentes mediante análisis de correlación y optimización de hiperparámetros, lo que no solo mejoró la precisión del modelo sino que también validó la importancia de características como la constante dieléctrica óptica mínima y los estados de oxidación en la predicción del bandgap. Además, el análisis de agrupamiento permitió identificar similitudes estructurales y electrónicas entre compuestos, mientras que la frecuencia de elementos, con el oxígeno como predominante, destacó la relevancia de los óxidos metálicos en este contexto.

La arquitectura del modelo de red neuronal se diseñó cuidadosamente, con una capa de entrada que integra las 10 características más relevantes, dos capas ocultas que procesan relaciones no lineales y una capa de salida para generar predicciones precisas. Las gráficas de evaluación evidenciaron que el modelo aún tiene potencial para mejorar mediante la incorporación de datos adicionales, lo que sugiere una escalabilidad prometedora.

En conclusión, este trabajo representa un avance significativo en la integración de la inteligencia artificial en el diseño de materiales. La capacidad predictiva demostrada no solo reduce costos y tiempos asociados a la síntesis y caracterización experimental, sino que también proporciona un marco adaptable para predecir otras propiedades materiales, contribuyendo al desarrollo acelerado de materiales funcionales para diversas aplicaciones tecnológicas.



1. References: L. Ge, Y. Ke, y X. Li, "Machine learning integrated photocatalysis: progress and challenges," Chemical Communications, vol. 59, no. 39, pp. 5795-5806, Apr. 2023. DOI: 10.1039/D3CC00899K.
 2. A. O. Ibhadow y P. Fitzpatrick, "Heterogeneous Photocatalysis: Recent Advances and Applications," Catalysts, vol. 3, no. 1, pp. 189-218, Mar. 2023. DOI: 10.3390/catal3010189.
 3. R. H. D. Santos, C. L. Oliveira, y J. C. Scorz, "Evaluation of XGBoost and other machine learning models for predicting the band gap of ZnO-based photocatalysts," Journal of Applied Physics, vol. 132, no. 14, pp. 145703, Oct. 2023. DOI: 10.1063/5.0065409.
 4. Y. Zhang, X. Zhang, y Y. Lu, "Advanced machine learning techniques for photocatalytic performance optimization," Materials Today Communications, vol. 31, pp. 103519, Jan. 2023. DOI: 10.1016/j.mtcomm.2022.103519.

