Modelización predictiva de las emisiones de CO_2 mediante regresión: Un enfoque hacia los Objetivos de Desarrollo Sostenible

Belén Rodríguez Llorente
Universidad Politécnica de Madrid (UPM)

Marzo 2025

Palabras clave: energías renovables, CO_2, crecimiento económico, machine learning.

1. Introducción

El presente estudio tiene como objetivo identificar el modelo de regresión más preciso para la estimación de emisiones de CO_2, permitiendo así analizar patrones y tendencias en la emisión de gases de efecto invernadero. A través del uso de técnicas de aprendizaje automático, se ha evaluado el desempeño de distintos enfoques predictivos con el fin de determinar el modelo más adecuado. Este trabajo se enmarca dentro de los Objetivos de Desarrollo Sostenible (ODS) propuestos por la ONU, contribuyendo específicamente al ODS 7 (Energía asequible y no contaminante), al proporcionar herramientas para evaluar el impacto de diferentes fuentes energéticas en las emisiones de CO_2; al ODS 9 (Industria, Innovación e Infraestructura), mediante el fomento de estrategias innovadoras para la reducción de emisiones en el sector industrial; y al ODS 13 (Acción por el clima), facilitando la toma de decisiones basada en datos para la mitigación del cambio climático.

1.1. Caracterización del problema

El problema se divide en varios bloques clave: recursos, razonamiento y objetivo. A continuación, se describen con mayor detalle, ya que el planteamiento y entendimiento del problema son fundamentales para abordarlo de manera efectiva.

1.1.1. Recursos

Conocimiento a priori: El cambio climático es una de las crisis ambientales más urgentes de la actualidad, impulsado en gran parte por las emisiones de gases de efecto invernadero (GEI), especialmente el dióxido de carbono (CO_2). Estas emisiones tienen su

origen en diversas actividades humanas, entre ellas la quema de combustibles fósiles, la industria y el transporte. La literatura científica ha evidenciado una estrecha relación entre el desarrollo económico, la demanda energética y la cantidad de emisiones generadas [1], lo que justifica la necesidad de desarrollar modelos predictivos para su análisis.

Datos y punto de partida: El estudio se basa en un conjunto de datos obtenido de Kaggle, disponible en Global Data on Sustainable Energy. Este dataset contiene 21 variables y 3,649 instancias relacionadas con factores energéticos, económicos y demográficos. No obstante, la calidad de los datos es un aspecto crucial, ya que se han identificado valores nulos y posibles inconsistencias en algunas columnas.

Para garantizar la precisión de los modelos desarrollados, se ha llevado a cabo un riguroso proceso de preprocesamiento, con especial atención a la imputación de valores faltantes. Se implementó una estrategia basada en la similitud entre países, aprovechando la correlación entre variables para seleccionar el país más adecuado con el fin de completar los datos de manera coherente. Sin embargo, esta estrategia introduce un sesgo inductivo, ya que asume que países con características similares presentan patrones energéticos y económicos comparables, lo que podría no ser válido en todos los casos. Además, el modelo aprende sobre un conjunto de datos preprocesado bajo estas suposiciones, lo que influye en su capacidad de generalización. En los casos en que una columna presentaba valores nulos en todas sus instancias, se recurrió a la identificación de países con características similares, mientras que, en columnas con valores nulos parciales, se utilizaron métricas de proximidad para determinar el país de referencia. De este modo, se minimizó la introducción de sesgos en la información utilizada para el entrenamiento del modelo.[1]

Herramientas Utilizadas: Para el desarrollo del análisis se emplearon diversas herramientas que facilitaron el procesamiento de datos y la construcción de modelos predictivos. Se utilizó Python como lenguaje de programación, junto con librerías especializadas como Pandas y NumPy para la manipulación de datos, Matplotlib y Seaborn para la visualización, y Scikit-learn, TensorFlow y Keras para el modelado. Como entorno de desarrollo se optó por Google Colab, que proporciona un espacio interactivo y accesible en la nube, complementado con Google Drive para la gestión eficiente del almacenamiento y acceso a los datos.

Distribución del Tiempo El proyecto se ha llevado a cabo en un período de dos semanas, lo que ha exigido una planificación estratégica para maximizar la eficiencia del análisis y la implementación del modelo. Este tiempo limitado ha permitido obtener resultados concretos, aunque también ha dejado abiertas diversas líneas de exploración futura.

1.1.2. Razonamiento matemático

El uso de modelos de aprendizaje automático, incluidas las redes neuronales artificiales, permite capturar relaciones complejas y no lineales entre variables, lo que las hace

[1]El cuaderno se encuentra disponible en el siguiente repositorio.

especialmente adecuadas para la predicción de emisiones de CO_2. En nuestro caso hemos hecho uso del KNN, del PMC y de las redes LSTM.

1.1.3. Objetivos

Este proyecto tiene como objetivo principal identificar el modelo de regresión más preciso para la estimación de emisiones de CO_2, desarrollando enfoques predictivos que respalden la toma de decisiones en políticas ambientales y energéticas. Al proporcionar herramientas basadas en datos, se busca apoyar estrategias efectivas de mitigación del cambio climático y promover un desarrollo sostenible. Este trabajo se alinea con los compromisos globales establecidos en los Objetivos de Desarrollo Sostenible (ODS), en particular con el ODS 7 (Energía asequible y no contaminante), ODS 9 (Industria, Innovación e Infraestructura) y ODS 13 (Acción por el clima), facilitando la implementación de soluciones fundamentadas en evidencia para reducir el impacto ambiental de las emisiones de gases de efecto invernadero.

2. Análisis de los datos

Para abordar el problema es fundamental comprender y analizar los datos que tenemos. En primer lugar, la Figura 1 muestra la evolución de las emisiones de CO_2 en los diez países con mayores emisiones a lo largo del tiempo. Se observa que China ha experimentado un crecimiento acelerado en sus emisiones, superando con creces a Estados Unidos, que históricamente ha sido uno de los principales emisores. India también ha incrementado sus emisiones de manera constante, reflejando así su crecimiento industrial y económico. Otros países como Alemania, Canadá y Japón presentan emisiones con leves reducciones, lo que sugiere esfuerzos en la implementación de fuentes más limpias.

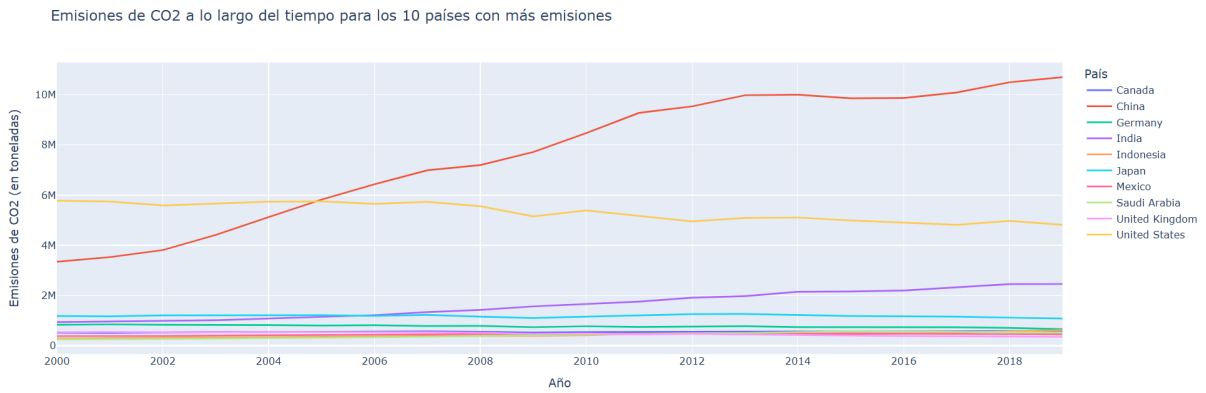


Figura 1: Emisiones de CO_2 a lo largo del tiempo para los 10 países con más emisiones.

Estos datos reflejan países altamente industrializados. Sin embargo, las tendencias también pueden indicar el impacto de políticas energéticas y la transición hacia fuentes más limpias.

La Figura 2 complementa este análisis al mostrar la evolución de la generación eléctrica en China y EE.UU., destacando el predominio de los combustibles fósiles en China y su creciente inversión en renovables, mientras que EE.UU. mantiene una mayor proporción de electricidad baja en carbono. Estas tendencias reflejan el desafío de reducir la dependencia del carbón y la necesidad de acelerar la transición hacia energías limpias en línea con la Agenda 2030, que busca incrementar la participación de fuentes renovables y reducir las emisiones globales de CO_2.

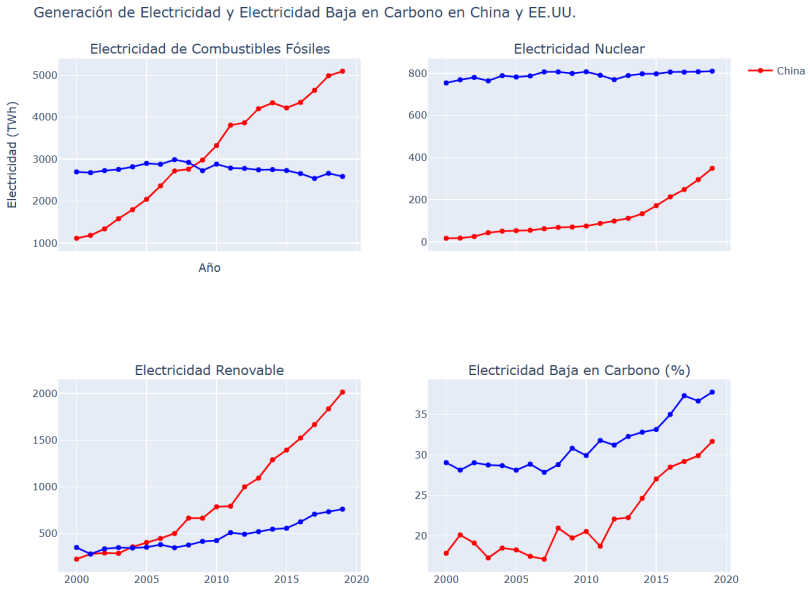


Figura 2: Generación de electricidad en China y EE.UU.

3. Metodología

Para la predicción de emisiones de CO_2, se emplean tres enfoques de modelado basados en aprendizaje automático: K-Nearest Neighbors (KNN), Perceptrón Multicapa (PMC) y redes LSTM. El procedimiento de resolución sigue las siguientes etapas:

3.1. Recolección y Preprocesamiento de Datos

Para el desarrollo del estudio, se utilizó un conjunto de datos históricos sobre emisiones de CO_2 junto con variables energéticas, económicas y geográficas.

3.2. Implementación de Modelos

Se emplean tres enfoques distintos para la predicción:

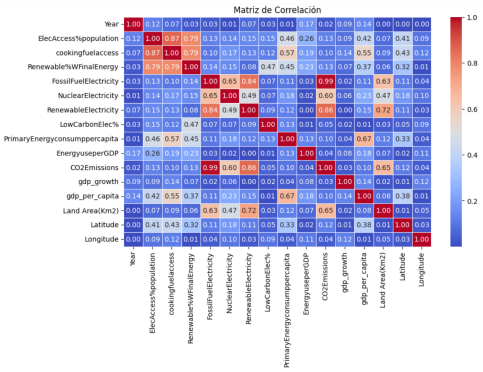


Figura 3: Matriz de correlación de Pearson.

Para garantizar la calidad de los datos y su adecuación al modelo, se aplicaron técnicas de normalización y escalado mediante la técnica *StandardScaler* para homogeneizar las magnitudes de las variables. Asimismo, se realizó un manejo exhaustivo de valores faltantes a través de métodos de imputación mediante interpolación, eliminación de registros en función de su impacto y sustitución por similitud con datos de países con características similares. Finalmente, se estableció un esquema de partición de datos con un 70 % para entrenamiento, 15 % para validación y 15 % para prueba, asegurando una evaluación efectiva del modelo.
Para escoger las mejores variables a predecir en los modelos, cogemos las más correladas con nuestro target.

KNN para Regresión

El algoritmo K-Nearest Neighbors (KNN) es un método de aprendizaje supervisado utilizado en este problema para regresión. Se basa en la idea de que los puntos con características similares tienden a agruparse en el espacio de características. KNN no construye un modelo explícito en la fase de entrenamiento, sino que almacena los datos y realiza los cálculos en el momento de la predicción.

Para determinar qué puntos son los vecinos más cercanos, KNN utiliza una métrica de distancia. La más común es la distancia Euclidiana, definida como:

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \tag{1}$$

donde p y q son puntos en un espacio de n dimensiones (en nuestro caso cogimos 11 features) y p_i, q_i representan sus respectivas coordenadas.

Otra distancia comúnmente usada es la de Manhattan:

$$d(p,q) = \sum_{i=1}^{n}|p_i - q_i| \tag{2}$$

De hecho, regularizando los parámetros el mejor modelo basándonos en el R^2 usa distancia Hamming.

En regresión, la predicción del valor de una nueva muestra x' se obtiene promediando los valores de sus k vecinos más cercanos:

$$\hat{y}(x') = \frac{1}{k}\sum_{i=1}^{K} y_i \tag{3}$$

donde y_i son los valores de los k vecinos más cercanos a x' en el conjunto de entrenamiento.

El valor de K influye en la precisión del modelo:

- Valores pequeños pueden hacer que el modelo sea sensible al ruido y tenga alta varianza.

- Valores grandes suavizan la predicción pero pueden perder precisión en la captura de patrones locales.

Una estrategia común para seleccionar el mejor K y los mejores hiperparámetros en general es evaluar los diferentes valores y elegir el que minimiza el error de predicción.

En este estudio, KNN se emplea para predecir las emisiones de CO_2 utilizando variables energéticas y económicas. Su capacidad para modelar relaciones no lineales lo convierte en una opción inicial antes de explorar modelos más avanzados como redes neuronales (PMC, LSTM).

Perceptrón Multicapa, PMC

Arquitectura del Perceptrón Multicapa (PMC) para Regresión

El Perceptrón Multicapa (PMC) es una red neuronal artificial perteneciente a la familia de redes feedforward, ampliamente utilizada para modelar relaciones complejas en datos numéricos.

Capa de Entrada La capa de entrada está compuesta por tantas neuronas como variables explicativas en el conjunto de datos. En este caso, 11.

Capas Ocultas Las capas ocultas permiten la extracción de patrones complejos mediante combinaciones de pesos sinápticos y funciones de activación no lineales. Se han evaluado distintas configuraciones arquitectónicas:

- $(32,)$: Una sóla capa oculta de 32 neuronas.
- $(64, 32)$: Primera capa oculta con 64 neuronas y segunda capa con 32 neuronas.
- $(128, 64, 32)$: Primera capa oculta con 128 neuronas, segunda con 64 neuronas y tercera con 32 neuronas.

Se han utilizado las funciones de activación ReLU y la tangente hiperbólica.

Capa de Salida Dado que el problema es de regresión, la capa de salida está compuesta por una única neurona que estima el valor continuo de emisiones de CO_2. En esta capa se emplea la función de activación identidad.

Función de Pérdida y Entrenamiento

Para minimizar la diferencia entre las emisiones de CO2 predichas y las reales, se utiliza el error cuadrático medio (MSE) como función de pérdida (viene predeterminada en la librería de Scikit-learn).

El entrenamiento se lleva a cabo mediante el algoritmo de retropropagación del error junto con el optimizador **Adam**, que combina Stochastic Gradient Descent (SGD) con momentum y adaptación de la tasa de aprendizaje.

Los criterios de detención del entrenamiento impuestos son el alcance de un número máximo de iteraciones predefinido y el estancamiento en la reducción del error de validación (early stopping).

Esta metodología proporciona una base sólida para la implementación y evaluación del PMC aplicado a la predicción de emisiones de CO_2, permitiendo analizar patrones y tendencias en la emisión de gases de efecto invernadero.

Long Short-Term Memory, LSTM

Arquitectura de la Red LSTM

Las redes de memoria a corto y largo plazo son un tipo de red neuronal recurrente diseñada para manejar secuencias de datos y problemas donde la dependencia temporal es relevante. A diferencia de las RNN tradicionales, las LSTM incorporan mecanismos de compuertas que regulan el flujo de información y permiten mitigar el problema del desvanecimiento del gradiente.

La arquitectura del modelo implementado en este estudio se compone de las siguientes capas:

Capa de entrada La entrada del modelo está diseñada para manejar secuencias de datos, lo cual es crucial para identificar patrones a través del tiempo en las series temporales de emisiones de CO_2, así como las variables relacionadas, como el acceso a electricidad, el consumo energético, la participación de energías renovables, entre otras.

Capas Ocultas Las capas ocultas del modelo LSTM son donde se capturan los patrones temporales en los datos. La LSTM es particularmente adecuada para este tipo de tarea porque puede recordar dependencias a largo plazo y manejar secuencias de datos con características complejas de temporalidad.

- **Primera capa LSTM (64 unidades):** Procesa las secuencias temporales y aprende relaciones a largo plazo entre los datos con dependencias no lineales.
- **Segunda capa LSTM (32 unidades):** Refina las representaciones de la primera capa, usando `return_sequences=False` para devolver una representación comprimida de la secuencia.

Esta reducción de la complejidad en la segunda capa permite que el modelo mantenga solo las características más relevantes y evite el sobreajuste, procesando más eficientemente las dependencias temporales.

Funciones de activación: Hemos usado tanh en las LSTM para introducir no linealidades y ReLU en las capas densas para aprender representaciones más complejas.

Capa de Salida La capa de salida es de tipo densa con una sola neurona, ya que estamos tratando con un problema de regresión donde se busca predecir un valor continuo: las emisiones de CO_2. La función de activación lineal es la adecuada, permitiendo así hacer predicciones continuas.

El modelo, por lo tanto, predice un valor escalar que representa la estimación de las emisiones de CO_2 para un dado conjunto de características.

La función de pérdida utilizada es el Error Cuadrático Medio (MSE).

El optimizador al igual que el PMC es Adam, lo que acelera la convergencia y mejora la estabilidad en el proceso de entrenamiento.

Los criterios de detención del entrenamiento impuestos son el alcance de un número máximo de iteraciones predefinido y el estancamiento en la reducción del error de validación (early stopping).

El modelo LSTM es ideal para predecir emisiones de CO_2 a partir de series temporales de datos energéticos, ya que captura las dependencias a largo plazo entre las variables a través de sus capas LSTM, permitiendo aprender patrones complejos en los datos. Este enfoque es particularmente adecuado para problemas con datos secuenciales que presentan relaciones no lineales, como las emisiones de CO_2, lo que hace de LSTM una excelente opción para este tipo de predicción.

3.3. Evaluación y comparación de modelos

En tareas de regresión, las métricas más comunes para evaluar la calidad del modelo son el R^2, el MSE (Error Cuadrático Medio) y el MAE (Error Absoluto Medio). El R^2 mide la proporción de la variabilidad de los datos que el modelo es capaz de explicar. Un valor de R^2 cercano a 1 indica que el modelo realiza una predicción precisa, mientras que valores cercanos a 0 sugieren que el modelo no ofrece mejoras sobre la media de los valores reales. Si R^2 es negativo, significa que el modelo es peor que simplemente predecir la media.

Por otro lado, el MSE calcula el promedio de los errores cuadrados y penaliza más los errores grandes, favoreciendo así un ajuste más preciso. El MAE, en cambio, calcula el promedio de los errores absolutos y es más robusto frente a valores atípicos, ya que no penaliza en exceso los errores grandes.

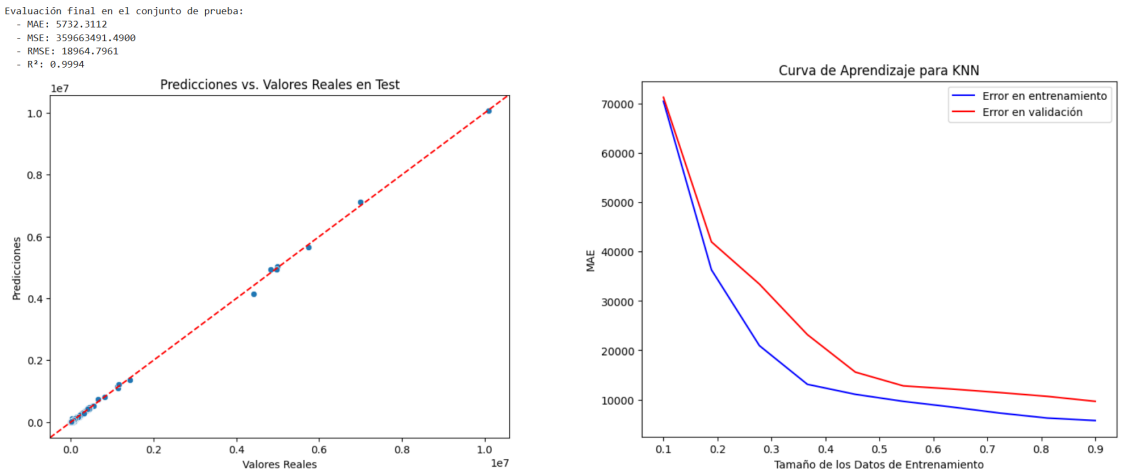
Aunque ambas métricas ofrecen información valiosa, en este caso, nos basaremos principalmente en el R^2 para seleccionar el mejor modelo, ya que proporciona una visión clara del rendimiento general del modelo en relación con la variabilidad de los datos.

4. Resultados

El modelo KNN aplicado para predecir las emisiones de CO_2 ha mostrado un rendimiento notablemente preciso, como se evidencia en las métricas obtenidas. Con un

coeficiente de determinación de 0.9994 en el conjunto de prueba, el modelo explica casi la totalidad de la variabilidad de los datos. A pesar del buen ajuste general, las métricas MAE (5732.31) y RMSE (18,964.80) indican desviaciones significativas en algunas predicciones, lo que sugiere la presencia de valores atípicos. Esto podría mejorarse ajustando hiperparámetros, explorando modelos más complejos o aplicando técnicas avanzadas de preprocesamiento.

Aunque la Figura 4(a) muestra una buena alineación entre predicciones y valores reales, la presencia de puntos dispersos sugiere dificultades en la estimación de valores extremos, lo que coincide con las métricas de error relativamente altas. La curva de aprendizaje en la Figura 4(b) indica un entrenamiento estable sin signos de sobreajuste o subajuste, pero el gráfico de residuos (Figura 5) revela errores más pronunciados en ciertos rangos. Esto refuerza la necesidad de mejorar el modelo mediante técnicas avanzadas de regresión para reducir la variabilidad en las predicciones y optimizar su precisión.



(a) Predicciones del KNN (b) Curva de aprendizaje del KNN

Figura 4: Predicciones y curva de aprendizaje del KNN

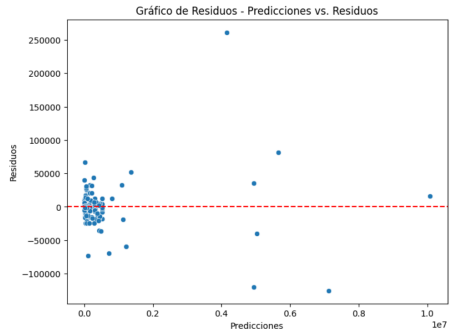


Figura 5: Gráfico de residuos del KNN

Dado que este estudio se enmarca dentro de los Objetivos de Desarrollo Sostenible (ODS), un modelo preciso de predicción de emisiones de CO_2 puede contribuir significativamente al análisis de tendencias y a la formulación de políticas enfocadas en la reducción del impacto ambiental. El uso de técnicas de aprendizaje automático representa una herramienta clave para la toma de decisiones basadas en datos, facilitando estrategias efectivas para mitigar el cambio climático.

El modelo PMC entrenado para predecir las emisiones de CO_2 ha mostrado un desempeño sólido, con un coeficiente de determinación de $R^2 = 0{,}9622$ en entrenamiento y $R^2 = 0{,}9587$ en validación, lo que sugiere una buena generalización. Además, en la evaluación final, se obtuvo un $R^2 = 0{,}9978$, con un MAE de 0.0280 y un RMSE de 0.0459,

Resultados de la búsqueda de hiperparámetros ordenados por menor diferencia en R^2:

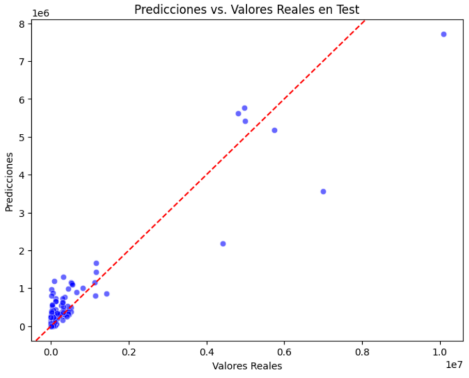
	Hidden Layers	Activation	Alpha	Learning Rate	Train MAE	Validation MAE	Train MSE	Validation MSE	Train RMSE	Validation RMSE	Train R^2	Validation R^2	R^2 Difference
26	(128, 64, 32)	relu	0.0010	constant	59519.575368	54314.912726	2.706996e+10	1.770148e+10	164529.520096	133046.908593	0.962217	0.958743	0.003475
27	(128, 64, 32)	relu	0.0010	adaptive	59519.575368	54314.912726	2.706996e+10	1.770148e+10	164529.520096	133046.908593	0.962217	0.958743	0.003475
24	(128, 64, 32)	relu	0.0001	constant	59547.865439	54355.258295	2.706098e+10	1.769875e+10	164502.207063	133036.646851	0.962230	0.958749	0.003481

Figura 6: Mejores hiperparámetros del PMC.

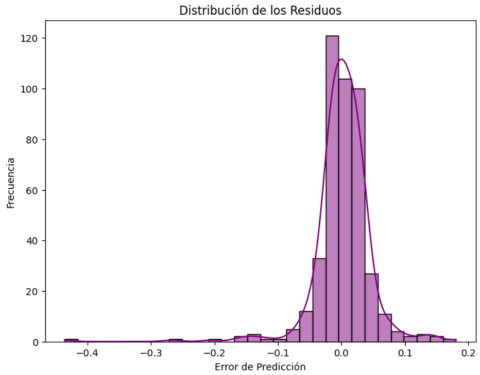
lo que indica un ajuste preciso..

En la Figura 7(a), la comparación entre valores reales y predicciones muestra una alineación cercana a la diagonal roja, pero con desviaciones en valores extremos. Asimismo, el gráfico de residuos en la Figura 7(b) refleja una distribución centrada en cero, aunque con cierta dispersión en los extremos, lo que sugiere inestabilidad en algunas predicciones.

Si bien el modelo logra capturar en gran medida la relación entre las variables, los errores observados en valores atípicos sugieren la necesidad de optimización. Ajustes en la arquitectura de la red, la normalización de datos o el uso de técnicas de regularización podrían mejorar su estabilidad y precisión.



(a) Predicciones del PMC.



(b) Residuos del PMC.

Figura 7: PMC

Finalmente, el modelo LSTM ha demostrado ser altamente efectivo para la predicción de emisiones de CO_2, logrando un coeficiente de determinación $R^2 = 0{,}9972$, lo que indica un ajuste casi perfecto entre los valores reales y las predicciones. Esto se puede observar en la Figura 8, donde la serie de datos reales y las predicciones muestran una gran similitud, reflejando la capacidad del modelo para capturar patrones temporales complejos.

En la gráfica de dispersión de predicciones frente a valores reales, también en la Figura 8, se aprecia una alineación clara con la diagonal roja, lo que sugiere que el modelo predice con gran precisión la mayoría de los valores. Sin embargo, se pueden notar ligeras desviaciones en valores extremos, lo que indica que el modelo podría beneficiarse de ajustes adicionales para mejorar su capacidad de generalización en situaciones menos comunes.

Los valores de error, representados en la Figura 9, muestran métricas sumamente bajas: MAE de 0.0317, MSE de 0.0164 y RMSE de 0.1281. Estos resultados refuerzan la idea de que el modelo LSTM es altamente preciso y presenta un error mínimo en la predicción

de emisiones. No obstante, sería recomendable realizar pruebas adicionales con distintos horizontes de predicción y técnicas de regularización para confirmar la estabilidad del modelo a largo plazo.

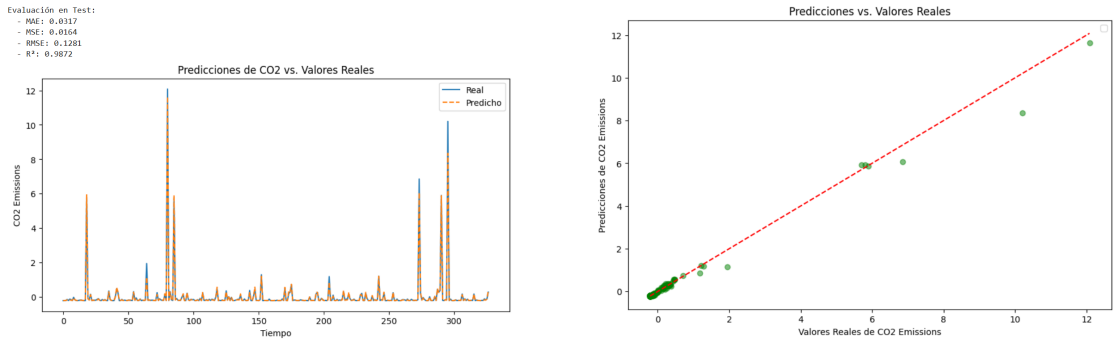


Figura 8: Predicciones de LSTM

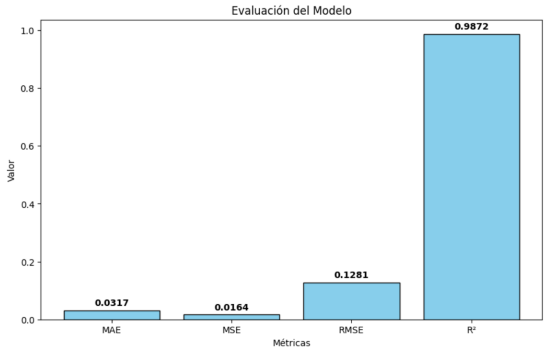


Figura 9: Métricas del LSTM.

En conclusión, el modelo LSTM se posiciona como una opción robusta para la estimación de emisiones de CO_2, con una capacidad destacada para capturar tendencias y patrones temporales. Aunque su desempeño es sobresaliente, sería útil explorar ajustes adicionales para mejorar su rendimiento en valores atípicos y validar su estabilidad en distintos conjuntos de datos.

5. Conclusiones y discusión

Los resultados obtenidos con el modelo LSTM demuestran un desempeño altamente preciso en la predicción de emisiones de CO_2, reflejado en un coeficiente de determinación $R^2 = 0{,}9972$. Este valor indica que el modelo es capaz de explicar casi la totalidad de la variabilidad en los datos, lo que lo convierte en una herramienta valiosa para el análisis de tendencias y la planificación de estrategias ambientales. Además, las métricas de error, como el MAE y el RMSE, son notablemente bajas, lo que sugiere que las predicciones presentan una desviación mínima respecto a los valores reales. Sin embargo, a pesar de este alto rendimiento, es fundamental considerar ciertas limitaciones. La sensibilidad a valores atípicos podría afectar la capacidad del modelo para generalizar en contextos con mayor variabilidad en los datos. Asimismo, dado que el modelo depende de datos históricos, su capacidad de adaptación a cambios estructurales, como nuevas regulaciones ambientales o avances tecnológicos en energías renovables, podría ser limitada.

Desde un enfoque ético, la aplicación de modelos de aprendizaje automático en la predicción de emisiones de CO_2 debe llevarse a cabo con responsabilidad. La precisión del modelo no solo tiene implicaciones en términos de análisis científico, sino también en la toma de decisiones políticas y económicas. Si los datos utilizados para el entrenamiento contienen sesgos o no representan adecuadamente la realidad global, las predicciones podrían conducir a interpretaciones erróneas y afectar negativamente el diseño de políticas públicas. Por ello, es crucial garantizar la transparencia en la recolección y uso de los datos, así como la evaluación continua del modelo en diferentes contextos.

Este trabajo está alineado con los Objetivos de Desarrollo Sostenible (ODS), contribuyendo especialmente a los ODS 7, 9 y 13. En primer lugar, el ODS 7, centrado en la energía asequible y no contaminante, se ve beneficiado por el análisis predictivo de emisiones de CO_2, ya que permite evaluar el impacto de distintas fuentes de energía y fomentar el uso de energías limpias. En relación con el ODS 9, que promueve la industria, la innovación y la infraestructura, la modelización de datos ayuda a que las industrias adopten estrategias más eficientes y sostenibles en términos de producción y consumo energético. Finalmente, este estudio tiene un impacto directo en el ODS 13, que aborda la acción por el clima, ya que el uso de modelos predictivos para analizar las emisiones de CO_2 proporciona herramientas valiosas para la toma de decisiones informadas en la lucha contra el cambio climático.

En conclusión, la inteligencia artificial y el aprendizaje automático tienen el potencial de ser aliados clave en la mitigación del cambio climático. Sin embargo, su aplicación debe ir acompañada de un uso ético de los datos, transparencia en los modelos (para evitar así sesgos inductivos innecesarios) y una evaluación constante de su impacto en la sociedad y el medio ambiente. La predicción precisa de emisiones de CO_2 no solo permite entender mejor la dinámica de los gases de efecto invernadero, sino que también ofrece una base sólida para el desarrollo de políticas efectivas en pro de la sostenibilidad global.

Referencias

[1] Mogota, A., & Djekonbe, D. (2022). Renewable energy and economic growth: The role of foreign direct investment in Sub-Saharan Africa. Renewable Energy, 12, 115-128.