



## Modelos neuronales profundos de la calidad del aire en la zona metropolitana de Tula, Hidalgo

### Deep numeral models of air quality in the Tula metropolitan area, Hidalgo

**Luis Felipe Barrera Cervantes**

Universidad Autónoma del Estado de Hidalgo, México

ba441259@uaeh.edu.mx

ORCID: 0009-0008-4641-6710

**Silvia Soledad Moreno Gutierrez**

Universidad Autónoma del Estado de Hidalgo, México

silviam@uaeh.edu.mx

ORCID: 0000-0002-8957-3707

doi: <https://doi.org/10.36825/RITI.13.29.010>

Recibido: Marzo 18, 2025

Aceptado: Mayo 12, 2025

**Resumen:** En las últimas décadas la creciente urbanización y desarrollo industrial han sido pilar para el desarrollo de zonas metropolitanas, lo que ha representado un incremento en la producción y emisión de gases de efecto invernadero. Esta situación ha contribuido a incrementar la contaminación del aire y ha dado lugar a una situación ambiental desfavorable que podría afectar a los ecosistemas, a la población y a la vida en general, por tanto, requiere la construcción de estrategias preventivas. Por lo anterior, se desarrollaron 2 modelos de inteligencia artificial para predecir la calidad del aire en la Zona Metropolitana de Tula (ZMT), mediante algoritmos de clasificación y regresión con redes neuronales de aprendizaje profundo. Se aplicó la metodología CRISP-DM, considerando la ZMT como productor del 90% de contaminación en Hidalgo, se realizó entendimiento de datos y su preparación, previo al modelado y validación de cada uno. Aplicando métricas de evaluación se obtuvo una coeficiente de determinación  $r^2$  de .99, pérdida de 0.088 para el modelo de regresión, para el de clasificación .93 de exactitud, precisión de .88 y  $F1$ -score de .81. Ambos modelos lograron representar el fenómeno de la calidad del aire y predecir de manera eficiente.

**Palabras clave:** *Calidad del Aire, Aprendizaje Profundo, Inteligencia Artificial, Contaminación Atmosférica, Monitorización Ambiental.*

**Abstract:** In recent decades, increasing urbanization and industrial development have been a cornerstone of metropolitan development, representing an increase in the production and emission of greenhouse gases. This situation has led to air pollution and an unfavorable environmental situation that could affect ecosystems, the population, and life in general. Therefore, preventive strategies are required. Therefore, two artificial intelligence models were developed to predict air quality in the Tula Metropolitan Area (TMA) using classification and regression algorithms with deep learning neural networks. The CRISP-DM methodology was applied, considering the TMA as the producer of 90% of pollution in Hidalgo. Data understanding and preparation were performed prior to modeling and validating each model. Applying evaluation metrics resulted in an  $r^2$  of .99, a loss of 0.088

for the regression model, a .93 accuracy for the classification model, a precision of .88, recall of .79 and *F1-score* of .81 Both models were able to represent the air quality phenomenon and predict it efficiently.

**Keywords:** *Air Quality, Deep Learning, Artificial Intelligence, Air Pollution, Environmental Monitoring.*

## 1. Introducción

Cuidar la salud es fundamental para asegurar la calidad de vida de las personas, en las últimas décadas la inteligencia artificial ha influido notablemente en este aspecto y ha contribuido al bienestar social [1]. No obstante, en diversas regiones del país la emisión de gases de efecto invernadero se ha vuelto un problema, sobre todo en zonas con fuerte actividad industrial como la ZMT donde los efectos desfavorables se acumulan, por lo que se considera una comunidad vulnerable.

La contaminación del aire es uno de los problemas ambientales más graves que enfrenta la humanidad, como consecuencia se estima que aproximadamente el 3% de las muertes por causas cardiopulmonares y el 5% de las muertes por cáncer de pulmón son atribuibles al material particulado [2]; además, la Organización Mundial de la Salud (OMS) ha declarado que el cambio climático representa la mayor amenaza para la salud [1]. Con el rápido desarrollo de la urbanización y la industrialización, diversos países en desarrollo sufren altos niveles de contaminación del aire, lo que ha generado una creciente preocupación tanto en gobiernos como en ciudadanos debido a su impacto en la salud humana y el desarrollo sostenible [3]. Según la OMS, la contaminación del aire es un “asesino silencioso” que provoca la muerte prematura de casi siete millones de personas cada año incluidos 600,000 niños [4].

Aunque en algunas regiones urbanas los niveles de contaminación por dióxido de azufre y carbono negro han disminuido, el aumento del tráfico vehicular ha intensificado la emisión de óxidos de nitrógeno, agravando el problema [5]. Los gases emitidos por fábricas y automóviles continúan elevando los niveles de contaminación atmosférica, lo que impacta negativamente en la calidad de vida. Por ello, predecir con precisión la calidad del aire es crucial para implementar medidas efectivas de control y mitigación [6].

El Índice de Calidad del Aire (ICA) se calcula a partir de seis contaminantes principales: CO (monóxido de carbono), NO<sub>2</sub> (dióxido de nitrógeno), PM<sub>10</sub>, PM<sub>1.0</sub>, CO<sub>2</sub> (dióxido de carbono) y O<sub>3</sub> (ozono) [7], sin embargo, la predicción del ICA requiere considerar factores adicionales como: humedad y presión atmosférica; además, el monitoreo efectivo debe clasificar si las condiciones son aptas o no para la habitabilidad humana [8]. Actualmente, los avances en inteligencia artificial, especialmente en el aprendizaje profundo, han abierto nuevas posibilidades para desarrollar modelos predictivos precisos y robustos. Los efectos de una mala calidad del aire impactan en la salud de las personas, incrementando el riesgo de molestias y enfermedades cardiovasculares y respiratorias, por ello, un sistema de monitoreo eficiente es de amplio beneficio para la salud humana y el control de la contaminación del aire [9].

El presente estudio tuvo como objetivo desarrollar dos modelos basados en técnicas de inteligencia artificial para predecir parámetros de calidad del aire utilizando redes neuronales mediante algoritmos de clasificación y regresión. Para ello, se empleó un conjunto de datos que corresponde a la ZMT, el cual incluye diversos atributos relacionados con la concentración de gases y variables climatológicas. El desarrollo del trabajo siguió la metodología CRISP-DM. Las contribuciones de este trabajo son el estudio comparativo de las técnicas de aprendizaje profundo aplicadas al análisis de la calidad del aire, así como la integración de modelos inteligentes con diferentes enfoques de forma complementaria. Entre los aportes destaca el desarrollo de modelos para predecir la calidad del aire a nivel local.

## 2. Estado del arte

En la última década, las técnicas de inteligencia artificial como el aprendizaje profundo, han sido integradas conjuntamente con las tecnologías de la información a entornos diversos [10] y han transformado la capacidad de modelar y predecir fenómenos complejos relacionados a la contaminación del aire, los métodos tradicionales como los basados en estadística lineal o modelos físicos de dispersión, han encontrado limitaciones significativas al enfrentar las interacciones complejas entre variables climáticas y contaminantes atmosféricos [11]. En contraste, los modelos de aprendizaje profundo han demostrado un rendimiento superior al procesar grandes volúmenes de datos y captar patrones no lineales propios de los sistemas ambientales [12].

### 2.1. Revisión de estudios previos

Con el auge actual de las redes profundas y la creciente demanda de información para anticipar problemas complejos, [11] emplearon redes de memoria a largo y corto plazo (LSTM por sus siglas en inglés) para predecir los niveles de PM<sub>2.5</sub> en Beijing, al entrenar el modelo con datos históricos sobre contaminantes y condiciones meteorológicas, lograron un  $r^2$  de 0.95 y una raíz del error cuadrático medio (RMSE por sus siglas en inglés) de 17.20, superando a los métodos de regresión lineal múltiple. Esta investigación subraya la capacidad de las redes neuronales recurrentes para captar dependencias temporales complejas, esenciales para predicciones precisas a largo plazo [13].

Sin embargo, la creciente innovación ha permitido el desarrollo de diversas opciones en las cuales se logra representar fenómenos complejos, en [14] propusieron un enfoque híbrido que combina redes neuronales convolucionales (CNN) para extraer características espaciales de mapas de contaminación, con redes ILSTM para modelar patrones temporales. Aplicaron este marco híbrido en Shijiazhuang city, Hebei Province, China, mejorando la precisión del ICA al integrar datos satelitales y de estaciones terrestres, alcanzando un MSE de 202.1923, un error absoluto medio (MAE por sus siglas en inglés) de 8.4134 y un  $r^2$  de 0.96. De igual manera, en [15] propusieron *Deep-AIR* teniendo el mismo enfoque de combinar CNN y LSTM para modelar la calidad del aire en ciudades metropolitanas, su enfoque aborda la compleja interacción entre la contaminación del aire y las dinámicas urbanas, logrando una precisión del 67.6% en estimaciones horarias de alta resolución en Hong Kong. Por otro lado, con el mismo enfoque en [16] presentaron un modelo híbrido que combina CNN y LSTM para pronósticos de contaminación del aire de alta resolución, su enfoque aprovecha las características espaciales y temporales de los datos de contaminación, logrando un error en entrenamiento de 15.7% y error en la validación de 27.1% siendo mejor en comparación con modelos de aprendizaje profundo existentes.

Con una perspectiva diferente, en [17] desarrollaron modelos LSTM para predecir la calidad del aire en Delhi enfocándose en el impacto del confinamiento por COVID-19, sus resultados indican que el modelo bidireccional LSTM multivariante ofreció las mejores predicciones destacando una media de error del 5.2% y un intervalo de confianza del 95% que refleja una mejora significativa en la precisión del modelo, evidenciando cambios notables en la calidad del aire durante y después del confinamiento.

Por otro lado, otra investigación propone la implementación de la LSTM para predecir niveles de contaminación de material particulado PM<sub>10</sub>, PM<sub>2.5</sub>, CO y NO<sub>2</sub> en la ciudad de Bogotá, utilizando tanto enfoques univariados como multivariados, e incorporando variables meteorológicas adicionales. Este trabajo logró un RMSE promedio de 5.38 para el modelo univariable y de 4.67 para el modelo multivariable, demostrando la efectividad de estas redes en la captura de patrones no lineales y la modelación de series temporales [18].

Por otra parte, en [19] desarrollaron modelos basados en atención que permiten capturar tanto relaciones locales como globales entre las variables climáticas, su método predice PM<sub>10</sub> y NO<sub>2</sub> utilizando datos de diversas ciudades asiáticas, superó a los modelos tradicionales al priorizar automáticamente las características más relevantes para cada contaminante alcanzando un error porcentual absoluto medio simétrico (SMAPE por sus siglas en inglés) de 0.2 en el desarrollo de su modelo.

Las investigaciones más recientes han adoptado enfoques híbridos, combinando técnicas avanzadas de aprendizaje profundo con mecanismos de atención, en [20] propusieron un modelo basado en un LSTM dual con mecanismos de atención espacio-temporal, además modelos de redes neuronales recurrentes (RNN) y *transformers* que permitió priorizar dinámicamente las características relevantes según el contexto geográfico y temporal, este modelo logró una precisión significativa en la predicción de la calidad del aire en la Cuenca, Ecuador obteniendo los mejores resultados en el modelo LSTM con un RMSE de 0.087, MAE de 0.056, un error porcentual absoluto medio (MAPE, por sus siglas en inglés) de 15.31%, un  $r^2$  de 0.701, superando a otros métodos tradicionales como SVR, XGBoost y regresión Ridge.

De manera similar, en [21] realizaron una revisión sistemática de la literatura sobre redes neuronales profundas aplicadas a la predicción de series temporales de calidad del aire, su análisis concluyó que los modelos híbridos que integran datos de múltiples fuentes (estaciones terrestres, datos satelitales y predicciones meteorológicas) son los más prometedores para abordar los desafíos actuales, alcanzando un  $r^2$  de 0.88 a 0.95 en predicciones a corto plazo superando en 20 a 25% a métodos tradicionales como ARIMA.

Posterior a esto, en [22] implementaron un enfoque de aprendizaje profundo espacio-temporal para la predicción regional de la calidad del aire, destacando la importancia de integrar datos de diferentes regiones para mejorar la generalización del modelo.

Sin embargo, en [23] propusieron un novedoso enfoque para la predicción del ICA en megaciudades, utilizando un modelo de aprendizaje multitarea espacio-temporal con mecanismos de atención, su propuesta logró alta precisión en la predicción de contaminantes como el SO<sub>2</sub> y el CO obteniendo MAE de 0.363 para SO<sub>2</sub> y 0.216 para CO y MAPE de 7.82 % y 2.14 %, respectivamente. Además, el modelo demostró una adecuada capacidad para capturar las tendencias de NO<sub>2</sub> y O<sub>3</sub>, a pesar de presentar una mayor variabilidad en los valores absolutos, lo que aporta información valiosa para la toma de decisiones en políticas de calidad del aire en entornos urbanos.

Por último, en [24] realizaron una revisión exhaustiva de 155 artículos relacionados, su análisis integra una amplia gama de enfoques, desde las redes neuronales artificiales (ANN) y máquinas de soporte vectorial (SVM, por sus siglas en inglés), hasta modelos avanzados de aprendizaje profundo como CNN, LSTM y arquitecturas híbridas. Entre los hallazgos del estudio se destaca que este modelo híbrido ofrece un mejor desempeño al combinar la captura patrones espaciales y temporales de los contaminantes, en algunos casos, logró reducir el MAE en un 15% en comparación con métodos tradicionales.

Además, los autores enfatizan que la integración de múltiples fuentes de datos, incluyendo mediciones de estaciones terrestres, datos satelitales y predicciones meteorológicas, mejora significativamente la precisión de las predicciones. En contraste, las principales limitaciones identificadas incluyen la dependencia de grandes volúmenes de datos para el entrenamiento y la necesidad de modelos más interpretables para facilitar su aplicación en la toma de decisiones. Finalmente, se recomienda que futuras investigaciones exploren técnicas para reducir el costo computacional y enfoques que prioricen la explicabilidad de los modelos.

## 2.2. Desafíos

Existen desafíos como la falta de datos homogéneos en áreas contaminadas y la necesidad de modelos que puedan ser aplicables a diferentes entornos geográficos [11]. Además, la implementación en tiempo real sigue siendo un obstáculo debido a la complejidad computacional de los modelos de aprendizaje profundo [12], a pesar de los avances logrados frente a otros enfoques tradicionales aún es necesario que futuras investigaciones se enfoquen en mejorar la capacidad de generalización de los modelos y optimizar su eficiencia.

## 3. Materiales y métodos

En este apartado se detalla la metodología empleada para el desarrollo de los modelos CRISP-DM (*Cross Industry Standard Process for Data Mining*). Este enfoque fue seleccionado por ser un marco sistemático y flexible que garantiza un flujo organizado desde la comprensión del problema hasta la generación de resultados replicables, enfatizando en el entendimiento del negocio.

### 3.1. Comprensión del negocio

El presente estudio busca analizar el impacto, así como contribuir al monitoreo ambiental mediante predicciones precisas que permitan identificar la calidad del aire y estimar valores específicos de contaminantes. A continuación se muestra la Tabla 1, que muestra las tareas efectuadas.

**Tabla 1.** Relación entre técnicas utilizadas y su predicción en el trabajo realizado.

Variable objetivo	Técnica	Determina
Calidad Aire	Clasificación	“Buena”, “Regular”, “Mala”
Calidad Aire	Regresión	Contaminante específico: “PM10”

Fuente: Elaboración propia.

Se consideró adecuado aplicar algoritmos de clasificación ya que se busca etiquetar la calidad del aire en clases previamente definidas, basadas en los valores de los contaminantes. Este enfoque es útil para alertas tempranas, recomendaciones de salud pública y toma de decisiones rápidas, ya que simplifica la interpretación del riesgo para la población vulnerable en la ZMT.

Los modelos de regresión son adecuados cuando el objetivo es estimar con precisión la concentración de un contaminante, PM10. En cuanto al análisis de partículas PM10, este se considera de amplia utilidad, ya que, según la revisión de la literatura, este contaminante es causante de cáncer pulmonar y muertes por enfermedades cardiopulmonares.

### 3.2. Comprensión de los Datos

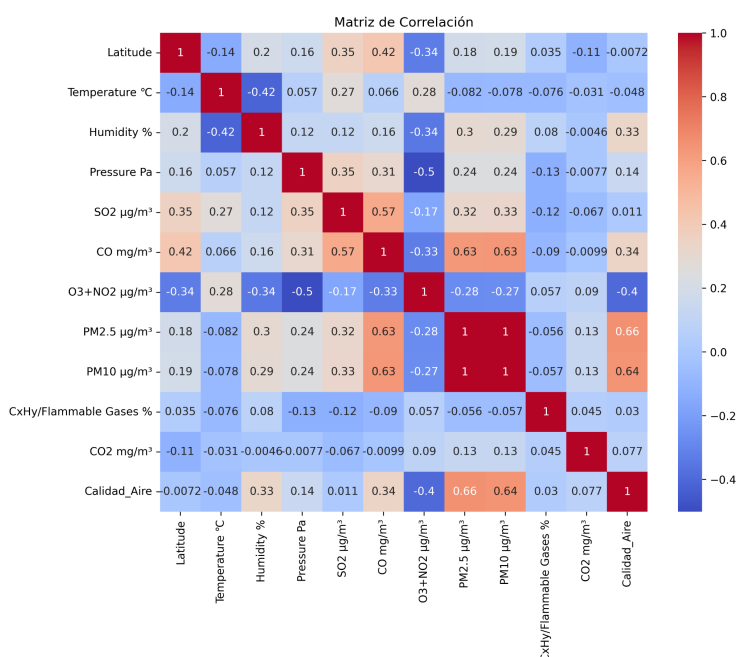
Los datos analizados fueron obtenidos de fuentes públicas, son consistentes con gases de efecto invernadero y meteorológicos. Ver Tabla 2.

**Tabla 2.** Datos utilizados, descripción y unidad de medida de cada característica y procesados por el modelo.

Tipo	Nombre	Descripción	Unidad de medida
Contaminante atmosférico	SO2	Dióxido de azufre	Miligramos por metro cúbico mg/m <sup>3</sup>
	CO	Monóxido de Carbono	Miligramos por metro cúbico mg/m <sup>3</sup>
	O3+NO2	Ozono + Dióxido de Nitrógeno	Miligramos por metro cúbico mg/m <sup>3</sup>
	PM10	Partículas diámetro <= 10 micrómetro	Micrómetros por metro cúbico
	PM2.5	Partículas diámetro <= 2.5 micrómetro	Micrómetros por metro cúbico
	CO2	Dióxido de Carbono	Miligramos por metro cúbico mg/m <sup>3</sup>
	CxHy/Flammable Gases	Gases inflamables	%
Variables meteorológicas	Humidity	Humedad	%
	Pressure	Presión en Pascales	Pascales
	Latitude	Latitud	Grados
	Temperature	Temperatura	Grados Celsius

Fuente: Elaboración propia.

Teniendo en cuenta las variables mencionadas se procedió a obtener la matriz de correlación para de esta manera identificar las características del conjunto de datos. Ver Figura 1. El conjunto de datos comprende un total de 17 columnas y 469 registros, cubriendo un rango temporal considerando el año 2022, los valores diarios se registraron con una frecuencia de cada 2 minutos por toma.



**Figura 1.** Matriz de correlaciones entre las características mismas.

### 3.3. Preparación de los datos

Para optimizar el rendimiento de los modelos, se implementaron las siguientes técnicas de preprocesamiento.

- a) Validación de datos existentes. Se hizo uso de la función *isnull().sum()* para confirmar que el conjunto de datos no contenga valores nulos ni inconsistencias que pudieran afectar el rendimiento de los modelos, esto evitó la necesidad de realizar procesos adicionales de limpieza y garantiza la calidad de los datos desde el inicio.
- b) Escalado de los datos. Para el escalado de los datos se aplicó la técnica *Standard Scaler* para ambos modelos, la cual consiste en transformar los datos y lograr una media de 0 y desviación estándar 1. Se aplicaron otras técnicas para realizar el escalado, no obstante, la que se menciona se consideró adecuada y se exponen los resultados en la sección correspondiente.
- c) Codificación de variables categóricas. Al buscar variables categóricas se obtuvo un resultado de seis variables las cuales son:
  - *Temperature* (°C).
  - *Humidity* (%).
  - *Pressure* (Pa).
  - *SO<sub>2</sub>* (µg/m<sup>3</sup>).
  - *O<sub>3</sub>+NO<sub>2</sub>* (µg/m<sup>3</sup>).
  - *CO<sub>2</sub>* (mg/m<sup>3</sup>).

Mediante el método *Label Encoding* las variables categóricas fueron transformadas a valores numéricos, esto fue necesario debido al formato original de los datos.

- d) Generación de características nuevas. Se generó una nueva columna llamada *Calidad\_Aire* basada en los valores de *PM2.5*, aplicando la siguiente condición:
  - Si *PM2.5* ≤ 12 se clasifica como Buena.
  - Si *PM2.5* > 12 y *PM2.5* < 45 se clasifica como Regular.
  - Si *PM2.5* > 45 se clasifica como Mala.

Este proceso se implementó en Python mediante el método *apply*, aplicado sobre la columna *PM2.5* µg/m<sup>3</sup>.

- e) División de los datos. Los datos se dividieron en conjuntos de entrenamiento y prueba para evaluar el desempeño de los modelos. El proceso se realizó mediante la función *train test split* de la librería *Scikit-learn* con los siguientes parámetros:
  - Características (*X*) incluyen las variables independientes mencionadas en la Tabla 2.
  - Etiqueta (*y*) la columna *Calidad\_Aire* previamente generada.
  - Tamaño del conjunto de prueba 30% de los datos *Test\_size = 32*.
  - Semilla aleatoria se estableció *random\_state = 42* para garantizar reproducibilidad.

Esto asegura que el conjunto de entrenamiento contenga el 70% de los datos, mientras que el 30% restante se utilizó para pruebas.

### 3.4. Modelado

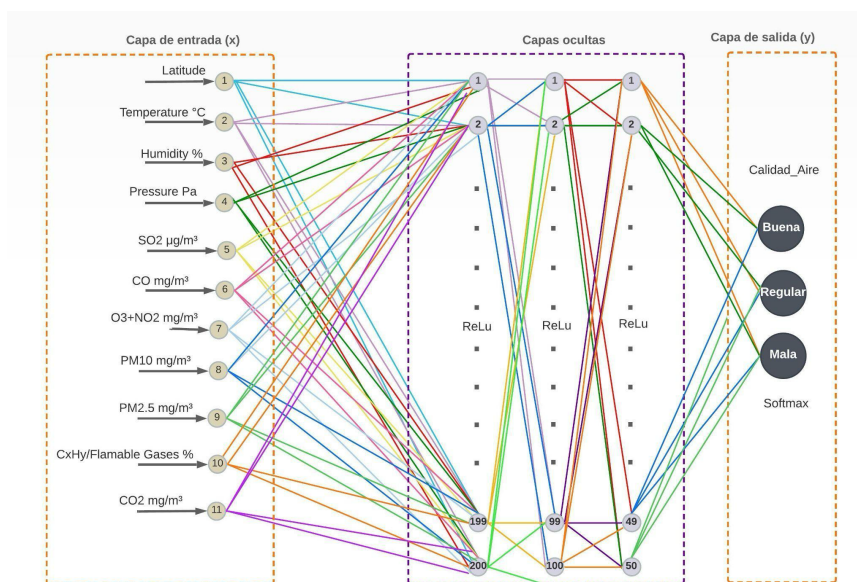
Con el objetivo de representar y predecir la calidad del aire en la ZMT se desarrollaron dos modelos utilizando técnicas de inteligencia artificial, cada uno se ajustó a los datos y objetivos del estudio, en este apartado se exponen las arquitecturas diseñadas.

#### 3.4.1. Modelo de clasificación: Perceptrón Multicapa (MLP)

Para este modelo se utilizó un perceptrón multicapa para clasificar la calidad del aire en los niveles previamente definidos, esta red neuronal permite capturar relaciones no lineales complejas entre las variables de entrada, lo que lo hace ser altamente usado en este tipo de tareas. A continuación se detalla la creación del modelo:

- El modelo se construyó para procesar un total de 11 variables de entrada, correspondientes a las características clave como *Latitude*, *CxHy Flammables gases*, *Humidity* (%), *Pressure* (Pa), *SO<sub>2</sub>* (µg/m<sup>3</sup>), *O<sub>3</sub>+NO<sub>2</sub>* (µg/m<sup>3</sup>), *CO<sub>2</sub>* (mg/m<sup>3</sup>).

- La primera capa oculta contiene 200 neuronas con una función de activación *ReLU*, adicional con *dropout* de 30%, al utilizar *dropout* se combate el sobre entrenamiento de la red, es decir, que vaya memorizando los datos, de manera que va utilizando subconjuntos alternativos en la red.
- La segunda capa oculta incluye 100 neuronas, con activación *ReLU*, adicional con *dropout* de 20%.
- La tercera capa oculta tiene 50 neuronas con activación *ReLU*, adicional con *dropout* de 20%.
- La capa de salida consiste en 3 neuronas con una activación *Softmax*, destinada a clasificar la calidad del aire en las categorías “Buena”, “Regular” o “Mala”, ver Figura 2.
- Se utilizó el optimizador Adam, por su rapidez en las redes neuronales profundas.



**Figura 2.** Arquitectura del modelo de clasificación durante el entrenamiento del modelo (ver Tabla 3).  
Fuente: Elaboración propia.

**Tabla 3.** Descripción de técnicas utilizadas en entrenamiento del modelo de clasificación.

Métrica	Aplicado	Fundamento	Ecuación
Función de pérdida	<i>Categorical Cross-Entropy</i>	Adecuada para problemas de clasificación multiclase, Ver Ecuación (1)	$CCE = - \sum_{i=1}^c Y_i \log(\hat{Y}_i) \quad (1)$
Métricas de evaluación	<i>Accuracy</i>	Para evaluar el desempeño del modelo, Ver Ecuación (2)	$Accuracy = \frac{VP + VN}{VP + FP + FN + FP} \quad dwsd(3)n \quad (2)$

Fuente: Elaboración propia.

Se utilizó la técnica de *Early Stopping* para evitar sobreajuste. El entrenamiento se configuró con las siguientes especificaciones:

- *Callbaks: Early Stopping*, monitoreando la pérdida en el conjunto de validación (*val\_loss*), con una paciencia de 10 épocas. Esto permitió restaurar los pesos del mejor modelo encontrado.
- Parámetros:
  - Épocas máximas: 200.
  - Tamaño del lote (*batch size*): 16.
  - Validación cruzada con el conjunto de prueba *validación\_data = X\_test, ytestcategorical*

El modelo fue entrenado utilizando los datos categóricos de entrenamiento (*X\_train, y\_train\_categorical*), y los resultados fueron almacenados en el objeto *history* para su posterior análisis.

### 3.4.2. Regla de aprendizaje (MLP)

Es relevante exponer la regla de aprendizaje del MLP, basado en algoritmo de *backpropagation*, que se encarga de ajustar los pesos de la red para minimizar el error en la salida. Para una red con  $L$  capas, donde la entrada es  $x$ , los pesos son  $w$  y la salida es  $y$ , la actualización de  $w$  se muestra en las Ecuaciones (3) y (4).

Propagación hacia adelante (*forward propagation*)

Cada neurona en la capa  $l$  calcula su salida con las Ecuaciones (3) y (4).

$$z^l = W^l a^{l-1} + b^l \quad (3)$$

$$a^l = f(z^l) \quad (4)$$

Donde:

$w^l$  son los pesos de la capa  $l$ .

$b^l$  es el sesgo.

$a^{l-1}$  es la activación de la capa anterior.

$f(z^l)$  es la función de activación, en este caso *ReLU* para capas ocultas y *Softmax* para capa de salida.

Cálculo de error en la salida

La pérdida se calcula aplicando CCE, que es la de mayor frecuencia en problemas de clasificación multiclase, ver Ecuación (5).

$$CCE = - \sum_{i=1}^c Y_i \log(\hat{Y}_i) \quad (5)$$

Donde:

$Y_i$  es la probabilidad real de la clase  $i$ .

$\hat{Y}_i$  es la probabilidad predicha por el modelo.

Propagación hacia atrás

Para minimizar el error se calcula el gradiente de la función de pérdida con respecto a  $w$ , aplicando la regla de cadena. La derivada de la pérdida respecto a la activación de la última capa se observa en las Ecuaciones (6), (7) y (8).

$$S^L = (\hat{Y} - Y) \quad (6)$$

$$\Delta w^L = \eta * \mathcal{E}^L * a^{L-1} \quad (7)$$

$$w^L = w^L - \Delta w^L \quad (8)$$

Donde:

$\eta$  Es la verdadera probabilidad de la clase  $i$ .

Propagación del error a capas anteriores

El error de la capa  $l$  se propaga hacia atrás y se representa en la Ecuación (9).

$$\mathcal{E}^l = (w^{l+1})^T \mathcal{E}^{l+1} \text{ ó } f'(z^l) \quad (9)$$

Donde:

$f'(z^l)$  Es la derivada de la función de activación.

Los  $w$  en cada capa se actualizan con las Ecuación (10) y (11).

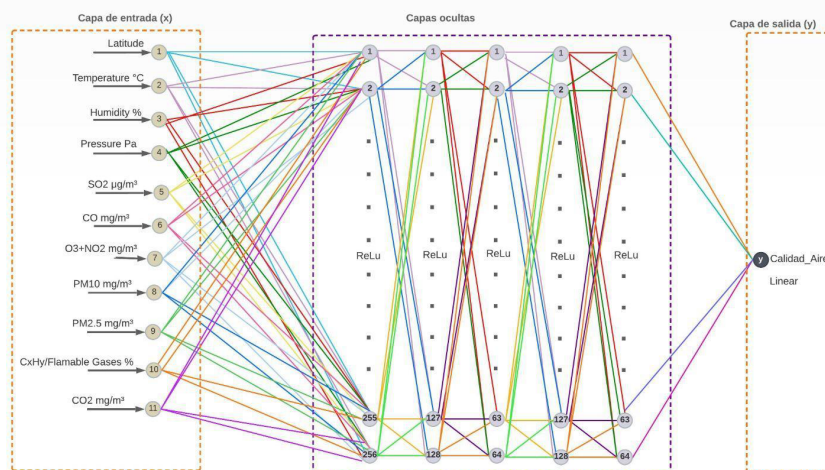
$$\Delta w^l = \eta * \mathcal{E}^l * a^{l-1} \quad (10)$$

$$w^l = w^l - \Delta w^l \tag{11}$$

### 3.4.2. Modelo de Regresión (ANN)

Para predecir concentraciones de contaminantes, se desarrolló un modelo de red neuronal artificial, este modelo trabaja con variables ambientales similares al modelo de clasificación pero enfocado en la predicción de valores continuos. A continuación, se detalla la arquitectura implementada:

- El modelo se construyó con las mismas características que el modelo de clasificación.
- Primera capa oculta: 264 neuronas con función de activación *ReLU*, seguida de un *dropout* del 20% para prevenir el sobreajuste.
- Segunda y cuarta capa oculta: 128 neuronas con función de activación *ReLU* y un *dropout* adicional del 20%.
- La tercera y quinta capa oculta tiene 64 neuronas con función de activación *ReLU* y un *dropout* de 20%.
- La capa de salida consiste en 1 neurona con una función de activación *Linear*, destinada a realizar predicciones continuas del material particulado encontrado en el aire. Ver Figura 3.
- Se utilizó el optimizador Adam, conocido por su eficiencia en la convergencia.
- Se aplicó la función *Early Stopping* que sirve para evitar el sobreajuste, este monitorea el *val loss* y tiene una *patience* de 10 épocas.



**Figura 3.** Arquitectura del modelo de regresión durante el entrenamiento del modelo (ver Tabla 4).  
Fuente: Elaboración propia.

En la Tabla 4 se exponen las técnicas utilizadas para la compilación y evaluación del modelo de regresión, en ella se describe la función de pérdida aplicada, así como la métrica utilizada para evaluar el rendimiento del modelo.

**Tabla 4.** Descripción de técnicas utilizadas en la del modelo de Regresión.

Métrica	Aplicado	Fundamento	Expresión matemática
Función de pérdida	MSE	Apropiada para problemas de regresión, ver Ecuación (12)	$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y})^2$ (12)
Métrica de evaluación	MAE	Para evaluar el desempeño del modelo ver Ecuación (13)	$MAE = \frac{1}{n} \sum_{i=1}^n  Y_i - \hat{Y} $ (13)

Fuente: Elaboración propia.

Para el entrenamiento del modelo consideró lo siguiente:

- Épocas máximas: 150.

- Tamaño del lote: 16.
- Validación: se utilizó el conjunto de prueba ( $validation\_data=(X\_test\_scaled, y\_test\_scaled)$ ).
- *Callbacks: Early Stopping* para monitorear la pérdida en validación ( $val\_loss$ ), con paciencia de 10 épocas para detener el entrenamiento si no se observaban mejoras.

Los resultados del entrenamiento fueron almacenados en el objeto `history` para su posterior análisis y evaluación.

### 3.5. Implementación y Herramientas

El desarrollo de los modelos se realizó utilizando:

- Lenguaje de programación: Python.
- Librerías y herramientas: TensorFlow, Keras, Scikit-learn, Pandas, NumPy y Matplotlib.
- Hardware: Computadora equipada de la marca Huawei Mate-Book D\_16 con procesador AMD Ryzen 5 4000H-Series, 16 GB de RAM, con 6 núcleos en el procesador y una GPU AMD Radeon Graphics para acelerar el entrenamiento.

## 4. Resultados

### 4.1. Evaluación del modelo para clasificación

El modelo de clasificación fue evaluado aplicando las métricas:

1. **Exactitud:** Mide la proporción de predicciones correctas entre el total de predicciones realizadas. Su fórmula está dada por la Ecuación (14), en la Tabla 5 se observa la exactitud obtenida por cada categoría en el modelo.

$$Exactitud = \frac{VP + VN}{VP + FP + FN + FP} \quad (14)$$

Donde:

*VP:* Verdaderos positivos

*VN:* Verdaderos negativos

*FP:* Falsos positivos

*FN:* Falsos negativos

**Tabla 5.** Exactitud por categoría de calidad del aire.

Etiqueta	Exactitud
Buena	0.84
Regular	0.80
Mala	1.0

Fuente: Elaboración propia.

El modelo obtuvo una exactitud de .93, lo que indica que el 93% de las veces que predijo una categoría el modelo acertó, esto refleja una alta confiabilidad en las predicciones.

2. **Recall:** también conocido como sensibilidad, es una métrica utilizada en la evaluación de modelos de clasificación. El *recall* mide la capacidad del modelo para identificar correctamente las muestras positivas utilizando la Ecuación (15), en la Tabla 6 se observa los resultados por cada categoría evaluada en el modelo.

$$Recall = 2 * \frac{VP}{VP + FN} \quad (15)$$

Donde:

*VP:* Verdaderos positivos.

*FN*: Falsos negativos.

**Tabla 6.** Recall por categoría de calidad del aire.

Etiqueta	Recall
Buena	1.0
Regular	0.40
Mala	0.98

Fuente: Elaboración propia.

El modelo obtuvo un *Recall* de .79 en promedio de las 3 categorías, lo que sugiere que el modelo fue capaz de identificar correctamente el 79% de los casos positivos reales.

3. *F1-score*: Esta métrica es útil en problemas con clases desbalanceadas, ya que combina la precisión y el *recall* en una sola métrica que se observa en la Ecuación (16), en la Tabla 7 se observa los resultados obtenidos por categoría en el modelo.

$$F1\ SCORE = 2 * \frac{Precisión * Recall}{Precisión + Recall} \tag{16}$$

Donde:

*Precisión*: Proporción de predicciones correctas entre todas las realizadas para una clase.

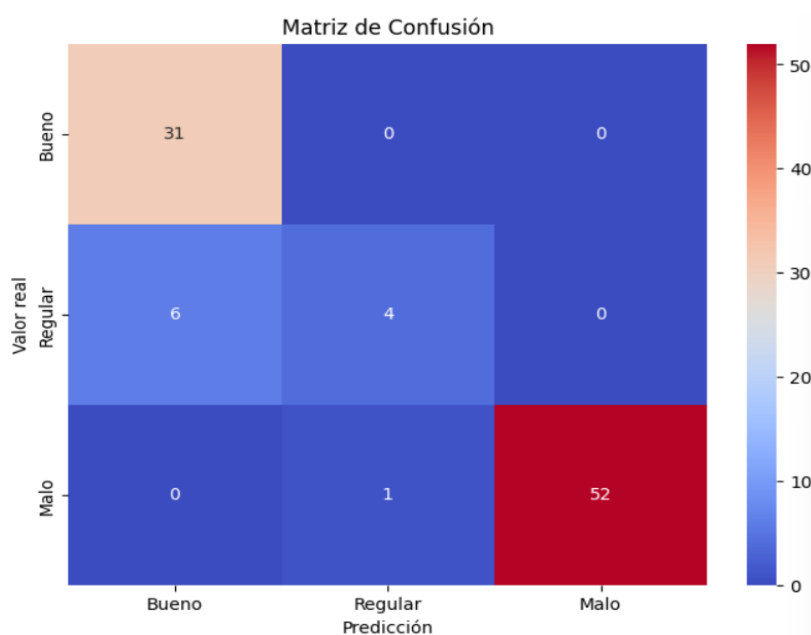
*Recall*: Proporción de positivos correctamente identificados.

**Tabla 7.** *F1-score* por categoría de calidad del aire.

Etiqueta	F1-score
Buena	0.91
Regular	0.53
Mala	0.99

Fuente: Elaboración propia.

El modelo obtuvo un *F1-score* de .81 en promedio de las 3 categorías lo que sugiere que el modelo tiene un buen equilibrio entre precisión y *recall*. Es una métrica especialmente útil cuando se requiere tener en cuenta tanto los falsos positivos como los falsos negativos. Se muestra la matriz de confusión obtenida en la Figura 4.



**Figura 4.** Matriz de confusión del modelo de clasificación.

Adicionalmente, al tener la matriz de confusión sugiere las métricas de especificidad y exactitud, las cuales se detallan.

4. *Especificidad*: también conocida como la Tasa de Verdaderos Negativos. Se trata de los casos negativos que el algoritmo ha clasificado correctamente. Expresa cuán bien puede el modelo detectar esa clase que se observa en la Ecuación (17).

$$\text{Especificidad} = \frac{VN}{VN + FP} \quad (17)$$

Donde:

*VN*: verdaderos negativos.

*FP*: falsos positivos.

La Tabla 8 presenta la especificidad de cada categoría en la clasificación, se observa alta capacidad para identificar correctamente las instancias que no pertenecen a una determinada categoría, alcanzando una especificidad del 96.43% en promedio de las 3 categorías, mostrando que la categoría “mala” puede ser clasificada bien en todos los casos.

**Tabla 8.** Especificidad por categoría de calidad del aire.

Etiqueta	Especificidad
Buena	0.90
Regular	0.98
Mala	1.0

Fuente: Elaboración propia.

5. *Precisión*: Es una métrica que mide qué tan precisas con las predicciones positivas de un modelo de clasificación y está dada por la Ecuación (18).

$$\text{precisión} = \frac{VP}{VP + FP} \quad (18)$$

Donde:

*VP*: verdaderos positivos.

*FP*: falsos positivos.

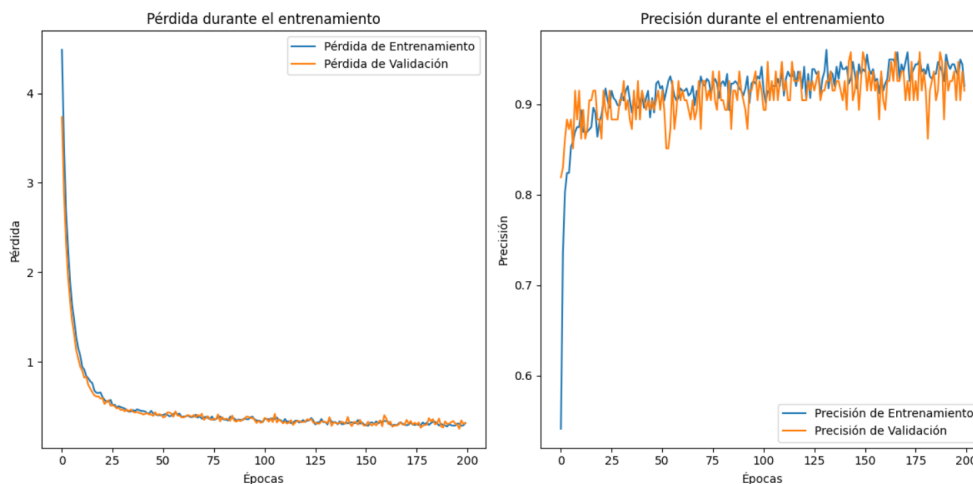
La Tabla 9 presenta la precisión de cada categoría, se observa alta capacidad para identificar correctamente las instancias que no pertenecen a una determinada categoría, alcanzando una precisión del 88% en promedio de las 3 categorías.

**Tabla 9.** Precisión por categoría de calidad del aire.

Etiqueta	Precisión
Buena	0.84
Regular	0.80
Mala	1.0

Fuente: Elaboración propia.

Así mismo para brindar una idea clara de la pérdida y la precisión durante el entrenamiento se realizó la siguiente gráfica ver Figura 5. El modelo muestra buen desempeño al mostrar que la pérdida es muy cercana a 0 y la precisión muy cercana a 1, así mismo se observan estables y se mantienen cercanas entre sí, sin señales evidentes de sobreajuste.



**Figura 5.** Vista gráfica de la pérdida y la precisión durante el entrenamiento.

#### 4.2 Evaluación del modelo para regresión

El modelo de regresión fue evaluado con las siguientes métricas:

1. MAE: Representa el promedio de las diferencias absolutas entre las predicciones y los valores reales, que está dado por la ecuación (19)

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (19)$$

Donde:

$n$ : número de ejemplo en el conjunto de datos.

$Y_i$ : Valor real para el  $i$ -ésimo ejemplo.

$\hat{Y}_i$ : Valor predicho por el modelo para el  $i$ -ésimo ejemplo.

$|Y_i - \hat{Y}_i|$ : Valor absoluto de la diferencia entre el valor real y el valor predicho.

El modelo alcanzó un MAE de 0.1187, sugiriendo que las predicciones del modelo están, en promedio, muy cerca de los valores reales, lo cual es lo más deseable.

2. El Coeficiente de Determinación ( $R^2$ ) es una métrica que indica la proporción de la variabilidad en los datos de respuesta que es explicada por el modelo de regresión. Es una medida de cuán bien los valores predichos aproximan los valores reales ver en la Ecuación (20).

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (20)$$

Donde:

$SS_{res}$ : Es la suma de los cuadrados de los residuos, es decir, las diferencias entre los valores observados y los valores predichos por el modelo.

$SS_{tot}$ : Es la suma total de los cuadrados, es decir, las diferencias entre los valores observados y la media de los valore observados.

Un  $R^2$  cercano a 1 indica que el modelo explica una gran proporción de la variabilidad en los datos de respuesta. En este caso, el  $R^2$  alcanzado por el modelo fue de 99%, lo que indica que el modelo es totalmente efectivo para predecir la variable objetivo en la cual se está trabajando a partir de las variables independientes seleccionadas.

3. Varianza explicada. Es una medida que cuantifica la proporción de la variabilidad total en los datos que es explicada por el modelo., es decir, mide qué tan dispersos están los datos alrededor de su media, ver Ecuación (21).

$$\text{Varianza explicada} = \frac{\lambda_i}{\sum_{j=1}^p \lambda_j} \quad (21)$$

Donde:

$\lambda_i$  Valor propio asociado al componente  $i$ .

$\sum_{j=1}^p \lambda_j$ : Suma de todos los valores propios.

La varianza explicada cercana a 1 sugiere que el modelo captura la mayoría de la variabilidad en los datos. En este caso, la varianza explicada alcanzada por el modelo es de 99%, indicando que el modelo es efectivo en representar los datos. La Figura 6 expone el comportamiento de la pérdida durante y la precisión durante el entrenamiento.

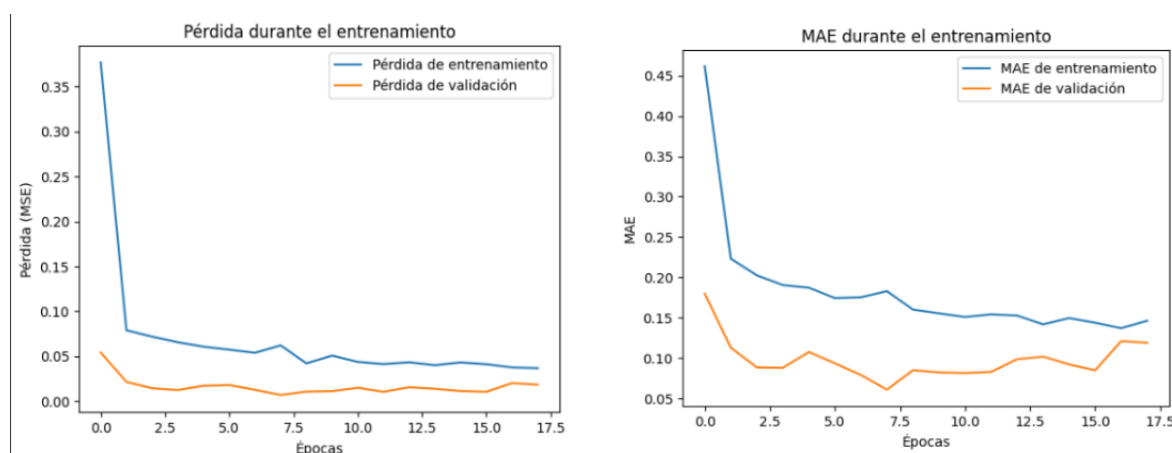


Figura 6. Vista gráfica de la pérdida y el error durante el entrenamiento.

Los resultados reflejan un buen desempeño del modelo con una reducción constante en la pérdida y del error absoluto medio durante el entrenamiento y la validación, además no se observan signos evidentes de sobreajuste lo que sugiere una adecuada capacidad de generalización.

## 5. Conclusiones

La investigación permitió evaluar el desempeño de modelos de aprendizaje profundo para estimar, predecir y clasificar la calidad del aire en la ZMT. Los resultados fueron satisfactorios, evidenciando un 99% de  $R^2$  en el modelo de regresión y un 93% de exactitud en el modelo de clasificación, lo que refleja un alto nivel de rendimiento y sugiere su potencial como herramientas confiables para apoyar la vigilancia ambiental y anticipar escenarios de riesgo en zonas vulnerables. Estos hallazgos coinciden con los objetivos de investigaciones previas reportadas en la literatura; sin embargo, los modelos propuestos lograron superar los resultados obtenidos anteriormente. Además, al estar basados en datos planos, presentan ventajas en términos de eficiencia y requerimientos computacionales.

A partir de lo anterior se plantea que investigaciones posteriores incorporen características diversas y enfoques que priorizan tanto la eficiencia como la administración de recursos, facilitando su aplicación en contextos diferentes.

Además como parte de trabajos a futuro se propone realizar predicción aplicando el modelo de regresión, y con base a sus resultados pasar el resultado por el modelo de clasificación, de esta manera con los rangos establecidos por el nivel de contaminación del ICA se obtendrá un modelo híbrido el cual permitirá mayor robustez con respecto a trabajos previos.

Un análisis más detallado del modelo de clasificación reveló que, en promedio, este logró una precisión del 88%, lo cual indica que la mayoría de las veces que predijo cada categoría acertó, por otro lado, el recall promedio fue del 79%, lo que significa que aunque el modelo reconoce bien las clases reales, aún tiene espacio para mejorar, especialmente en la categoría "Regular", donde su desempeño fue limitado, el *F1-score* con un valor de 0.81, mostró que el modelo mantiene un equilibrio entre lo que predice correctamente y lo que deja de identificar.

Otro aspecto importante es la especificidad promedio, que alcanzó el 96.4%, demostrando que el modelo evita las falsas alarmas, es decir, no clasifica erróneamente otras clases como si fueran, por ejemplo, "Mala", esta última categoría, de hecho, mostró los mejores resultados: precisión perfecta, recall del 98%, *F1-score* de 0.99 y especificidad del 100%. Esto resulta valioso desde el punto de vista de salud pública y prevención ya que permite detectar con precisión los eventos de mala calidad del aire.

Como proyección a futuro, se propone enriquecer los modelos incorporando variables adicionales, como factores meteorológicos o patrones de actividad industrial para incrementar su capacidad de generalización, así mismo se plantea el diseño de un modelo híbrido: primero una estimación de concentración mediante regresión y posteriormente se calificaría el nivel de calidad del aire usando los rangos del ICA, esta combinación permitiría obtener resultados precisos y útiles integrando la potencialidad de ambos enfoques.

## 6. Referencias

- [1] World Health Organization: WHO. (2023). *Cambio climático*. <https://www.who.int/es/news-room/fact-sheets/detail/climate-change-and-health>
- [2] Benito Lozano, M., Rodríguez González, S. (2023). La imperiosa necesidad de programar la asistencia a urgencias, atendiendo al cambio climático y las crisis de calidad del aire. *Revista Española de Urgencias y Emergencias*, 2, 194-197. <https://digital.csic.es/handle/10261/339047>
- [3] Li, X., Peng, L., Hu, Y., Shao, J., Chi, T. (2016). Deep learning architecture for air quality predictions. *Environmental Science and Pollution Research*, 23, 22408–22417. <https://doi.org/10.1007/s11356-016-7812-9>
- [4] South China Morning Post. (2021). *Air pollution kills 7 million a year, says WHO as it tightens guidelines*. <https://www.scmp.com/news/world/europe/article/3149735/air-pollution-kills-7-million-year-says-who-it-tightens>
- [5] Wang, J., Li, J., Wang, X., Wang, J., Huang, M. (2020). Air quality prediction using CT-LSTM. *Neural Computing And Applications*, 33 (10), 4779-4792. <https://doi.org/10.1007/s00521-020-05535-w>
- [6] Chen, H., Guan, M., Li, H. (2021). Air Quality Prediction Based on Integrated Dual LSTM Model. *IEEE Access*, 9, 93285-93297. <https://doi.org/10.1109/ACCESS.2021.3093430>
- [7] Zou, X., Zhao, J., Zhao, D., Sun, B., He, Y., Fuentes, S. (2021). Air Quality Prediction Based on a Spatiotemporal Attention Mechanism. *Mobile Information Systems*, 2021 (1), 1-12. <https://doi.org/10.1155/2021/6630944>
- [8] Hamami, F., Dahlan, I. A. (2022). *Air Quality Classification in Urban Environment using Machine Learning Approach*. IOP Conference Series Earth and Environmental Science. <https://doi.org/10.1088/1755-1315/986/1/012004>
- [9] Chakma, A., Vizena, B., Cao, T., Lin, J., Zhang, J. (2017). *Image-based air quality analysis using deep convolutional neural network*. IEEE International Conference on Image Processing (ICIP). Beijing, China. <http://doi.org/10.1109/ICIP.2017.8297023>
- [10] Ruiz Muñoz, G. F., Vasco Delgado, J. C. (2025). Integración de las tecnologías de la información y la comunicación (TIC) e inteligencia artificial (IA) en la formación docente. *Revista de Investigación en Tecnologías de la Información*, 13 (29), 60-70. <https://doi.org/10.36825/RITI.13.29.006>
- [11] Zhang, M., Wu, D., Xue, R. (2021). Hourly prediction of PM2.5 concentration in Beijing based on Bi-LSTM neural network. *Multimedia Tools Applications*, 80, 24455–24468. <https://doi.org/10.1007/s11042-021-10852-w>
- [12] Qiu, W., Zhou, Y., He, H., Wang, B., Mu, G., Zhou, M., Ma, J., Zhang, Z., Cao, L., Xiao, L., Yuan, J., Chen, W. (2021). Short-term effects of air pollution on liver function among urban adults in China. *Atmospheric Environment*, 245. <https://doi.org/10.1016/j.atmosenv.2020.118011>

- [13] Zhao, F., Liang, Z., Zhang, Q., Seng, D., Chen, X. (2021). Research on PM2.5 Spatiotemporal Forecasting Model Based on LSTM Neural Network. *Computational Intelligence and Neuroscience*, 2021 (1), 1-10. <https://doi.org/10.1155/2021/1616806>
- [14] Wang, J., Li, X., Jin, L., Li, J., Sun, Q., Wang, H. (2022). An air quality index prediction model based on CNN-ILSTM. *Scientific Reports*, 12 (8373), 1-16. <https://doi.org/10.1038/s41598-022-12355-6>
- [15] Han, Y., Zhang, Q., Li, V. O. K., Lam, J. C. K. (2021). Deep-AIR: A Hybrid CNN-LSTM Framework for Air Quality Modeling in Metropolitan Cities. *arXiv preprint*. <https://arxiv.org/abs/2103.14587>
- [16] Zhang, Q., Lam, J. C. K., Li, V. O. K., Han, Y. (2020). Deep-AIR: A Hybrid CNN-LSTM Framework for Fine-Grained Air Pollution Forecast. *arXiv preprint*. <https://arxiv.org/abs/2001.11957>
- [17] Tiwari, A., Gupta, R., Chandra, R. (2021). Delhi air quality prediction using LSTM deep learning models with a focus on COVID-19 lockdown. *arXiv preprint*. <https://arxiv.org/abs/2102.10551>
- [18] Sarmiento Sánchez, C. A. (2024). Predicción de Contaminantes Atmosféricos en Bogotá utilizando Redes LSTM. *Ingenio Tecnológico. Revista de Ingeniería*, 6 (E051), 1-26. <https://ingenio.frlp.utn.edu.ar/index.php/ingenio/article/view/108>
- [19] Mengara Mengara, A. G., Park, E., Jang, J., Yoo, Y. (2022). Attention-Based Distributed Deep Learning Model for Air Quality Forecasting. *Sustainability*, 14 (6), 1-34. <https://doi.org/10.3390/su14063269>
- [20] Buestán Andrade, P. A., Carrión Zamora, P. E., Chamba Lara, A. E., Pazmiño Piedra, J. P. (2025). A comprehensive evaluation of ai techniques for air quality index prediction: RNNs and transformers. *Ingenius*, (33), 60-75. <https://doi.org/10.17163/ings.n33.2025.06>
- [21] Zaini, N. A., Ean, L. W., Ahmed, A. N., Malek, M. A. (2022). A systematic literature review of deep learning neural network for time series air quality forecasting. *Environmental Science and Pollution Research*, 29 (4), 4958-4990. <https://doi.org/10.1007/s11356-021-17442-1>
- [22] Abirami, S., Chitra, P. (2021). Regional air quality forecasting using spatiotemporal deep learning. *Journal of Cleaner Production*, 283. <https://doi.org/10.1016/j.jclepro.2020.125341>
- [23] Khan, H., Tso, J., Nguyen, N., Kaushal, N., Malhotra, A., Rehman, N. (2024). Novel Approach for Predicting the Air Quality Index of Megacities through Attention-Enhanced Deep Multitask Spatiotemporal Learning. *arXiv preprint*. <https://arxiv.org/abs/2407.11283>
- [24] Méndez, M., Merayo, M. G., Núñez, M. (2023). Machine learning algorithms to forecast air quality: a survey. *Artificial Intelligence Review*, 56, 10031–10066. <https://doi.org/10.1007/s10462-023-10424-4>