



ARQUITECTURA DE IoT PARA EL MONITOREO DE EMISIONES DE GASES CONTAMINANTES DE VEHÍCULOS Y SU VALIDACIÓN A TRAVÉS DE MACHINE LEARNING

IoT ARCHITECTURE FOR VEHICLE POLLUTANT GAS EMISSION MONITORING AND VALIDATION THROUGH MACHINE LEARNING

Washington Torres Guin^{1,*} , José Sánchez Aquino¹ ,
 Samuel Bustos Gaibor¹ , Marjorie Coronel Suárez¹ 

Recibido: 16-11-2023, Recibido tras revisión: 12-12-2023, Aceptado: 16-02-2024, Publicado: 01-07-2024

Resumen

Este estudio propone una arquitectura IoT para el monitoreo de emisiones de gases contaminantes en vehículos, en respuesta a la creciente preocupación por la contaminación del aire y el calentamiento global. La arquitectura se basa en un nodo equipado con sensores DHT22, MQ9 y MQ135 para capturar la temperatura, humedad y emisiones de gases, mismo que se comunica de manera efectiva a través de la red LTE para enviar los datos a la plataforma *ThingSpeak*. Se lleva a cabo un análisis de los niveles de contaminación de CO_2 , CO y CH_4 mediante los datos recopilados. Estos datos se validan mediante la revisión técnica de un vehículo de prueba. Posterior, se entrena una red neuronal artificial (ANN) utilizando una base de datos específica de emisiones de CO_2 de vehículos en Canadá, como resultado se obtiene un R^2 alto de 99,2 % y los valores de RMSE y MSE bajos, esto indican que el modelo está haciendo predicciones precisas y se ajusta bien a los datos de entrenamiento. La ANN tiene como objetivo predecir las emisiones de CO_2 y verificar los datos de CO_2 provenientes de la red IoT. La arquitectura demuestra su capacidad para el monitoreo en tiempo real y su potencial para contribuir a la reducción de la contaminación.

Palabras clave: contaminación vehicular, gases contaminantes, IoT, LTE, sensores

Abstract

This study proposes an IoT architecture for monitoring emissions of polluting gases in vehicles, in response to the growing concern about air pollution and global warming. The architecture is based on a node equipped with DHT22, MQ9, and MQ135 sensors to capture temperature, humidity, and gas emissions, respectively. This node effectively communicates through the LTE network to send the data to the ThingSpeak platform. An analysis of CO_2 , CO , and CH_4 pollution levels is conducted using the collected data. These data are validated through the technical review of a test vehicle. Subsequently, an Artificial Neural Network (ANN) is trained using a specific database of CO_2 emissions from vehicles in Canada. As a result, a high R^2 of 99.2 % is achieved, along with low values of RMSE and MSE, indicating that the model is making accurate predictions and fits well to the training data. The ANN aims to predict CO_2 emissions and verify CO_2 data from the IoT network. The architecture demonstrates its capability for real-time monitoring and its potential to contribute to pollution reduction.

Keywords: Vehicle pollution, pollutant gases, IoT, LTE, sensors

^{1,*}Facultad de Sistemas y Telecomunicaciones, Universidad Estatal Península de Santa Elena, Santa Elena, Ecuador.
 Autor para correspondencia ✉: wtorres@upse.edu.ec.

Forma sugerida de citación: Torres Guin, W.; Sánchez Aquino, J.; Bustos Gaibor, S. y Coronel Suárez, M. "Arquitectura de IoT para el monitoreo de emisiones de gases contaminantes de vehículos y su validación a través de Machine Learning," *Ingenius, Revista de Ciencia y Tecnología*, N.º 32, pp. 9-17, 2024. DOI: <https://doi.org/10.17163/ings.n32.2024.01>.

1. Introducción

La urbanización acelerada y el aumento en la demanda de transporte han agravado los problemas ambientales asociados al transporte por carretera [1]. A pesar de la disminución temporal de las emisiones en 2020 debido a la pandemia de COVID-19, el calentamiento global persiste con el aumento de la contaminación atmosférica [2]. Los retos del cambio climático trascienden fronteras y afectan a nivel global, ya que la emisión de gases de efecto invernadero como metano (CH_4), hidrofluorocarbonos (HFC), dióxido de carbono (CO_2) altera el equilibrio entre la Tierra y la atmósfera. En particular, las emisiones de CO_2 se han convertido en un factor crucial en el agravamiento del calentamiento global, liberando aproximadamente ocho mil millones de toneladas anualmente debido a la quema de combustibles fósiles en el transporte, la calefacción y la producción de energía [3]. Estas emisiones, que provienen principalmente de la actividad industrial y vehicular, siendo los automóviles de pasajeros responsables del 75 % de las emisiones de dióxido de carbono [4].

Los vehículos pesados, que incluyen autobuses y camiones, contribuyen con el 25 % aproximadamente de las emisiones totales en carreteras, y existe una tendencia preocupante que podría aumentar este porcentaje si no se toman medidas adecuadas. A pesar de las regulaciones cada vez más estrictas que buscan supervisar la eficiencia del combustible y controlar las emisiones de gases de efecto invernadero, se ha evidenciado un notable aumento en la cantidad de automóviles en circulación. Este incremento se ha traducido en un aumento significativo en la cantidad de kilómetros recorridos por estos vehículos, lo que ha acentuado aún más su contribución a las emisiones de contaminantes atmosféricos [5].

La vigilancia de la contaminación en el ámbito vehicular resulta fundamental por varias razones [6,7]. Ofrece datos esenciales acerca del origen y la intensidad de la contaminación del aire en entornos de transporte público, privado, carga, entre otros, información crucial para la formulación de políticas y medidas destinadas a reducir la exposición y mejorar la calidad general del aire. Asimismo, este seguimiento contribuye a identificar enfoques efectivos para abordar el problema, como la implementación de combustibles más limpios y la adopción de tecnologías avanzadas, desempeñando un papel esencial en la búsqueda de soluciones sostenibles y en la promoción de entornos urbanos más saludables tanto para los viajeros como para la población en general.

Donde, el desafío radica en la carencia de datos precisos y confiables acerca de la calidad del aire en el transporte vehicular, lo que impulsa la necesidad de diseñar y probar un sistema de medición móvil capaz de suplir esta deficiencia [8]. La captura, el procesamiento y el análisis de datos sobre los niveles de contamina-

ción en el transporte urbano resultan cruciales para mejorar la comprensión de las fuentes de contaminación del aire en este entorno, y a su vez, incentivar la elaboración de políticas e intervenciones con miras a abordar este desafío.

A continuación, se presenta una revisión de las propuestas IoT para el monitoreo de emisiones. Senthilkumar *et al.* [9] describen un sistema integrado en el cual los sensores recopilan datos de calidad del aire y los transmiten a nodos de niebla. Moses [10] propone un esquema basado en la nube para monitorear la calidad del aire utilizando sensores que calculan niveles de contaminantes como NO_x , CO, O_3 , PM10, PM2.5 y SO_2 , junto con datos ambientales como humedad y temperatura. Los datos recopilados se actualizan en la nube a través de una infraestructura de Lora Gateway y nodos Lora. Se utiliza el análisis de series temporales junto con modelos de regresión de vectores de soporte y redes neuronales de perceptrón multicapa para predecir la concentración de partículas contaminantes. Behal y Singh [11] emplean el método ANFIS para prever la calidad del aire en función del nivel de contaminantes y el índice de calidad del aire modificado (m-AQI). Donde, un modelo de regresión de vectores de soporte se utiliza para pronosticar valores, lo que implica determinar una línea de mejor ajuste que sea resistente a valores atípicos.

Shetty *et al.* [12] aplican métodos de IoT para supervisar las tasas de emisiones vehiculares y utilizan datos reales a nivel mundial para pronosticar los niveles de monóxido de carbono. Wei *et al.* [13] emplean el monitoreo vehicular para proporcionar a los propietarios detalles sobre la contaminación actual en su ubicación y la tasa de emisiones de sus vehículos, todo ello con el respaldo de técnicas de aprendizaje automático para predecir la contaminación basándose en datos históricos y actuales recopilados por sensores. Mumtaz *et al.* [14] dan una solución que combina sensores de IoT avanzados con capacidades de aprendizaje automático para monitorear y predecir la calidad del aire en entornos interiores, lo que permite la medición de diversos contaminantes y en el trabajo de Mohamed [15] se utiliza una red de sensores IoT para detectar ocho tipos de contaminantes y, mediante técnicas de aprendizaje automático, se logra una alta precisión en la clasificación de la calidad del aire interior, con un 99,1 % de exactitud.

Por lo tanto, en un mundo cada vez más consciente de la importancia de la sostenibilidad y la reducción de la contaminación ambiental, la monitorización de emisiones de gases contaminantes en vehículos se ha vuelto un desafío crítico. La creciente urbanización y la expansión del parque automotor han planteado la necesidad imperiosa de controlar y mitigar la contaminación del aire, tanto para preservar la calidad del entorno como para salvaguardar la salud pública. En este contexto, la Internet de las Cosas (IoT) emerge

como una herramienta poderosa que permite la recopilación y análisis de datos en tiempo real, ofreciendo la capacidad de supervisar y gestionar las emisiones vehiculares de manera eficiente y efectiva.

Este proyecto se centra en el desarrollo de una arquitectura de IoT para la monitorización de emisiones de gases contaminantes en vehículos, respaldada por técnicas de Machine Learning. Esta arquitectura permitirá la recopilación en tiempo real de datos precisos sobre las emisiones de vehículos en funcionamiento, su posterior análisis y la validación de estos datos a través de algoritmos de aprendizaje automático. La implementación de esta propuesta reviste una importancia crítica por diversas razones.

- **Control de emision.** El monitoreo en tiempo real de las emisiones de vehículos ofrece la posibilidad de identificar y gestionar de manera proactiva fuentes de contaminación, lo que es esencial para alcanzar objetivos de calidad del aire y reducir el impacto ambiental.
- **Tecnología y sostenibilidad.** La combinación de IoT y Machine Learning representa un enfoque tecnológico avanzado que refuerza la sostenibilidad y demuestra cómo la innovación puede ser un motor de cambio positivo en la sociedad.

2. Materiales y métodos

En la siguiente sección se expone la arquitectura propuesta, así como los elementos que la componen, destinados a la construcción del nodo sensor para la adquisición de gases en un vehículo de prueba, específicamente un Kia Sportage del año 2012.

2.1. Arquitectura IoT propuesta

En esta arquitectura de Internet de las Cosas (IoT) (Figura 1), se ha diseñado un sistema completo para la monitorización de emisiones de gases contaminantes en vehículos, se toma como referencia las propuestas de [16, 17]. La arquitectura consta de varias capas que trabajan en conjunto para recopilar, almacenar y validar los datos de emisiones.

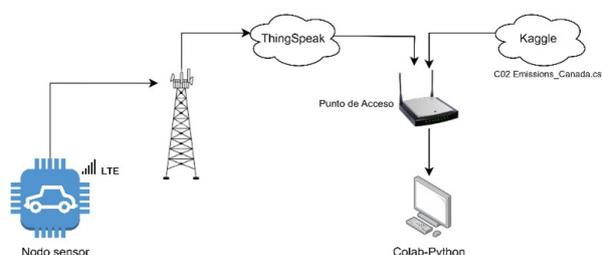


Figura 1. Arquitectura IoT para el monitoreo de gases contaminantes

A continuación, se describe la arquitectura basándose en sus capas y su funcionalidad.

- **Capa de percepción.** En la capa de percepción, se usa un nodo equipado con un sensor de temperatura DHT22, así como sensores MQ-135 y MQ-9 para la detección de gases. Este nodo es responsable de la captura de datos en tiempo real relacionados con la temperatura, humedad y la concentración de gases contaminantes en el aire.
- **Capa de red.** La capa de red o comunicación es esencial para transmitir los datos capturados por el nodo sensor. Se ha optado por utilizar la red LTE como medio de comunicación, lo que permite una conectividad confiable y amplia cobertura para el nodo sensor y posible escalabilidad de la red.
- **Capa de aplicación.** En la capa de aplicación se ha integrado la plataforma IoT ThingSpeak para almacenar y gestionar los datos recopilados por los sensores. ThingSpeak ofrece una interfaz amigable y permite el almacenamiento seguro de los datos a través del protocolo HTTP, lo que facilita su acceso y análisis posterior.
- **Análisis de los datos.** Una vez que los datos se encuentran en ThingSpeak, se procede a realizar análisis más detallados para verificar la distribución de los datos de emisiones de dióxido de carbono (CO_2), monóxido de carbono (CO), metano (CH_4). Se utiliza la herramienta de Colab, una plataforma de colaboración para la programación en Python.
- **Validación de datos.** Con el objetivo de garantizar la precisión de los datos capturados por los sensores, se lleva a cabo una verificación exhaustiva que incluye pruebas fundamentadas en la revisión técnica de un vehículo de prueba. Además, se realiza una regresión utilizando una base de datos de emisiones de CO_2 de origen canadiense como punto de referencia para comprobar los datos que llegan de la plataforma IoT. Este procedimiento posibilita la comparación, validación y predicción.

2.2. Diseño del nodo sensor

El diseño del nodo sensor (Figura 2) se ha configurado de manera integral utilizando el controlador LILYGO@ TTGO T-Call V1.4, que incorpora una variedad de características esenciales. Este dispositivo ofrece conectividad a través de la red LTE porque dispone de un módulo SIM800L, además de contar con la potencia del ESP32 para capacidades inalámbricas, como wifi y Bluetooth. Así mismo, el nodo sensor está equipado con un sistema de posicionamiento GPS incorporado,

lo que permite la geocalización precisa de las mediciones. En cuanto a los sensores utilizados, se han seleccionado cuidadosamente para garantizar la obtención de datos precisos. El sensor DHT22, conocido por su fiabilidad, se encarga de medir la temperatura y la humedad del entorno. Por otro lado, los sensores MQ135 y MQ9 se utilizan para detectar la concentración de CO, CO_2 y metano (CH_4). Cabe destacar que los sensores DHT22 se conectan a uno de los puertos digitales disponibles en el controlador, mientras que los sensores MQ135 y MQ9 se conectan a puertos analógicos, lo que proporciona una interfaz versátil para la adquisición de datos. Esta configuración integral del nodo sensor permite una medición precisa y la recopilación de datos clave para la monitorización de emisiones de gases contaminantes en vehículos.

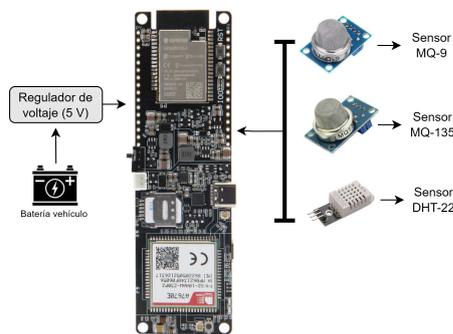


Figura 2. Nodo Sensor IoT

El sensor DHT22 opera con un voltaje de 3,3–5 VCC y es capaz de medir la humedad relativa del aire en un rango de 0–99 % RH, con una precisión de ± 2 % (a 25 °C) y una resolución de hasta 0,1 %. La temperatura se mide en el rango de -40 – 80 °C, con una precisión de $\pm 0,5$ °C y una resolución de 0,1 °C. El sensor tiene una velocidad de actualización de 1 Hz (reporta cada 1 segundo) y utiliza el protocolo Wire. Para su funcionamiento, se emplea la biblioteca DHT.h.

El sensor MQ-135 opera mediante un mecanismo de detección específico que involucra interacciones con gases, lo que resulta en variaciones en su resistencia eléctrica. A pesar de que el sensor no discrimina gases de forma intrínseca, se puede calibrar y configurar para detectar gases particulares en función de sus patrones de respuesta únicos. Al ajustar con precisión los parámetros del sensor y aplicar técnicas avanzadas de procesamiento de señales, es factible distinguir diversos gases y sus concentraciones, permitiendo al sensor diferenciar con eficacia entre gases como CO y CO_2 . De acuerdo con la información técnica proporcionada por el sensor, se establece que la resistencia de carga del sensor es de 20,1 k Ω , mientras que la resistencia del sensor en condiciones de aire limpio es de 10 k Ω . Con estos datos se presenta en la Figura 3 la curva de calibración del sensor con la ecuación de modelo, donde,

Ro se define como una constante que representa la resistencia del sensor en respuesta a una concentración de 0,4 mg/L, mientras que Rs denota la resistencia del sensor en otro contexto.

El sensor MQ-9 se utilizó para determinar la concentración de metano (CH_4). Los datos de salida analógica del sensor se envían a la entrada analógica del controlador. La conversión de los datos recibidos del convertidor analógico-digital (ADC) se realiza en un rango de 0 a 3,3 V. Las características y especificaciones, como la carga y la resistencia al aire fresco, son similares a las del MQ-135, ya que ambos sensores son del mismo fabricante y comparten los mismos valores, por ende, se aplican las mismas configuraciones.

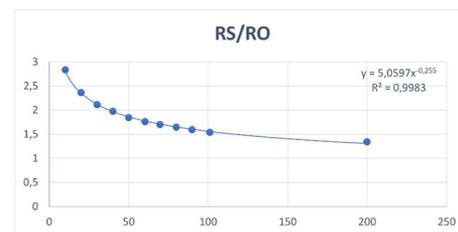


Figura 3. Curva de calibración del sensor MQ-135

2.3. Implementación del nodo sensor al vehículo

El nodo sensor IoT se posiciona siguiendo las indicaciones proporcionadas en la Figura 4, lo que permite una colocación estratégica de los sensores directamente en la salida del tubo de escape del vehículo. El controlador, por su parte, se sitúa de manera segura dentro del vehículo, estableciendo una conexión directa con la computadora a bordo. A través de esta conexión, se logra la visualización en tiempo real de los datos capturados por el sensor. Además, el sensor se conecta a la plataforma ThingSpeak mediante la tecnología LTE, utilizando el protocolo HTTP para la transmisión y almacenamiento eficiente de los datos recopilados. Este diseño integral permite un monitoreo efectivo de las emisiones de gases contaminantes en vehículos, proporcionando información valiosa y en tiempo real. Para la parte de energía el dispositivo se conecta a la batería del vehículo.

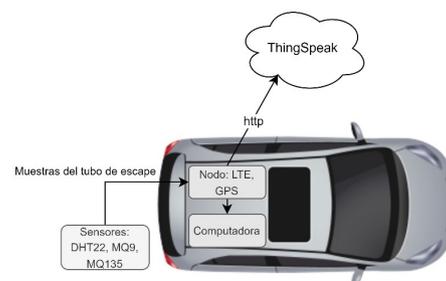


Figura 4. Disposición del nodo sensor IoT en el vehículo

En la Figura 5 se presenta claramente el nodo sensor instalado en el vehículo para la recolección de muestras.

En la Figura 6 se evidencian los sensores que se colocaron en el tubo de escape del vehículo.

Es importante destacar que el GPS se emplea con el propósito de visualizar la posición precisa del vehículo y obtener un detallado registro de la trayectoria que ha seguido.

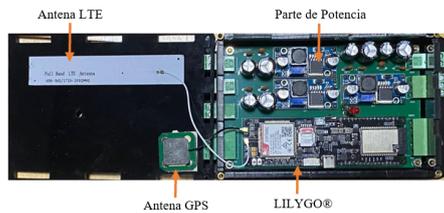


Figura 5. Nodo sensor IoT instalado

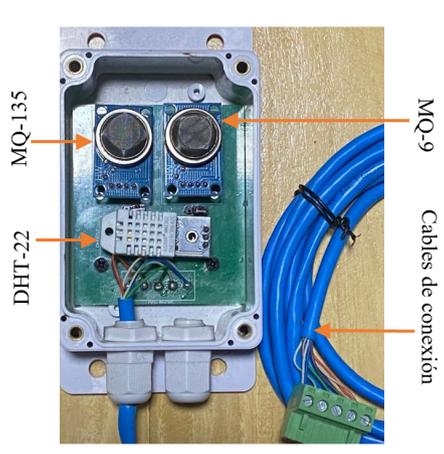


Figura 6. Sensores para la toma de datos

3. Resultados y discusión

En primer lugar, se presentan los resultados obtenidos a partir de las 50 muestras recopiladas del tubo de escape con el motor en ralentí y con ciertos cambios en las revoluciones de este para identificar la variabilidad en la toma de los datos. Estas muestras revelan los niveles de contaminantes, específicamente CH_4 (metano), CO_2 y CO. Los datos proporcionados por los sensores se expresan en partículas por millón (ppm), y los resultados ofrecen una visión detallada de las emisiones.

Al analizar los resultados, se observa que la concentración de CH_4 (Figura 7) varía aproximadamente entre 45 y 65 ppm, reflejando la distribución de los datos. Además, en la Figura 8 se destaca que el CO muestra oscilaciones en el rango de 26 000 a 38 000 ppm y CO_2 se evidencia que sus niveles oscilan entre 121 000 y 140 000 ppm.

Es relevante señalar que estos valores se sitúan dentro de los rangos especificados, tal como se menciona en el estudio [17]. Estos hallazgos subrayan la importancia de monitorear de forma detallada las emisiones vehiculares, proporcionando una comprensión clara de la variabilidad de los contaminantes del motor. Como dato importante las muestras se toman de 3 a 4 minutos aproximadamente esto debido al tiempo que lleva el controlador en procesar y publicar estos datos en la plataforma ThingSpeak.

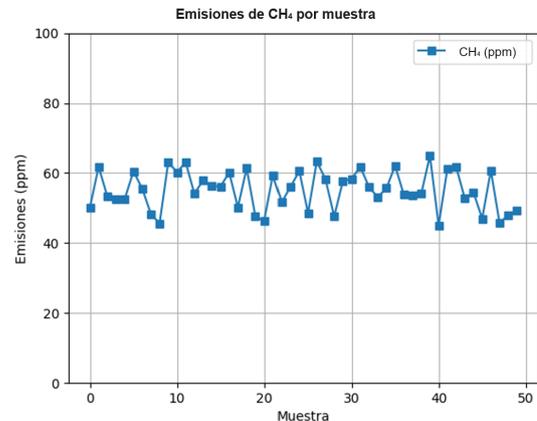


Figura 7. Gráfica de los datos de CH_4

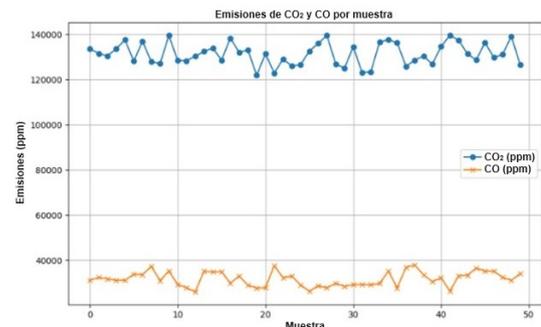


Figura 8. Gráfica de los datos de CO y CO_2

La representación gráfica de los datos de temperatura y humedad no fue realizada porque estos parámetros no varían bruscamente, según la posición del sensor en el tubo de escape. Durante las mediciones la temperatura demostró oscilaciones mínimas, manteniéndose en un rango constante de aproximadamente 30 a 37 °C. Simultáneamente, la humedad exhibió una estabilidad similar, fluctuando entre el 80 % y el 99 %.

La validación de la captura de datos se llevó a cabo mediante la realización de una exhaustiva revisión técnica vehicular en un SUV Kia Sportage del año 2012, utilizado como vehículo de prueba. Como resultado de esta evaluación, se obtuvieron concentraciones de CO_2 en el rango de 12 % a 14 % con el motor en ralentí, y concentraciones de CO en el rango de 2,6 a 3,8 %. Según lo indicado por Segura [18], se puede aplicar este conocimiento para estimar las emisiones en

ppm, situando el CO_2 en su nivel máximo estimado de 140 000 ppm y el CO en 38 000 ppm.

Esta información respalda la confiabilidad de los datos recopilados, ya que los resultados de la revisión técnica vehicular coinciden con las mediciones obtenidas por los sensores. La calibración precisa del sensor para la captura de estos gases refuerza la precisión de las estimaciones, demostrando que la concentración de CH_4 también es precisa y confiable.

Para evaluar el nivel de contaminación del vehículo, centrándonos en el CO_2 como una variable de referencia crucial dada su significativa contribución a las emisiones vehiculares, se realiza una conversión con el objetivo de obtener las unidades de g/km que representan los gramos de contaminación por kilómetro recorrido. La relevancia de este dato radica en su verificación a través de la fuente o ficha técnica de vehículo que estipula un nivel de contaminación de 158 g/km de CO_2 en entornos urbanos.

Por lo tanto, para determinar cuánto dióxido de carbono (CO_2) se libera al recorrer un kilómetro, si el vehículo marco un porcentaje de 14 % de concentración, ese porcentaje significa que el 14 % del volumen del gas en el tubo de escape del vehículo es CO_2 , y el 86 % restante es otro tipo de gases de escape, como nitrógeno, oxígeno e hidrocarburos no quemados (HC) entre otros.

Para calcular la cantidad de CO_2 , se necesita considerar varios factores, como la eficiencia del vehículo, la cantidad de combustible que quema y la cantidad de CO_2 que se produce por litro de combustible quemado.

Eficiencia del vehículo. Este parámetro se toma de la ficha técnica del vehículo, donde se evidencia que tiene un consumo de combustible de 5,7 litros por cada 100 kilómetros.

Cantidad de CO_2 por litro de combustible. La cantidad de CO_2 producida por la quema de un litro de gasolina varía según la composición exacta del combustible, para este caso esta alrededor de 2,8 kg de CO_2 por litro de gasolina, con estos parámetros se procede aplicar la siguiente ecuación (1) para encontrar el CO_2 en g/km.

$$\begin{aligned} CO_2 \left(\frac{g}{km} \right) &= C_{l/km} * CO_{2kg/l} * 100_{g/kg} \\ CO_2 \left(\frac{g}{km} \right) &= 0.057_{l/km} * 2,8_{\frac{kg}{l}} * 100_{\frac{g}{kg}} \\ &= 159,6 \text{ g/km} \end{aligned} \quad (1)$$

Este cálculo demuestra una aproximación casi exacta con la ficha técnica del vehículo donde se tiene una estimación de 158 g/km de contaminación de CO_2 .

Con este dato de conversión establecido, se procede a desarrollar un modelo de red neuronal artificial (ANN) con el objetivo de prever las posibles emisiones de gases contaminantes en el parque automotor. Para este propósito, se usa una base de datos de Kaggle [19] que alberga información sobre emisiones de gases contaminantes vehiculares en Canadá. La elección de esta

base de datos se fundamenta en la presencia de información sobre el mismo vehículo utilizado en nuestras pruebas de sensores. De esta manera, se crea un modelo de aprendizaje que se adapta a nuestro escenario, permitiendo la predicción del nivel de contaminación de CO_2 .

La base de datos contiene parámetros cruciales, como la marca (Make), modelo (Model), clase de vehículo (Vehicle Class), tamaño del motor (Engine Size), número de cilindros (Cylinders), tipo de transmisión (Transmission), tipo de combustible (Fuel Type), consumo de combustible en ciudad (Fuel Consumption City (l/100 km)), consumo de combustible en avenidas (Fuel Consumption Hwy (l/100 km)), así como las emisiones de CO_2 (CO_2 Emissions) medidas en gramos por kilómetro (g/km). Durante el análisis, se realiza una correlación entre estos parámetros y las emisiones de CO_2 para identificar las variables con una relación más significativa. Aquellas variables que presentan una correlación notable se seleccionan para el entrenamiento del modelo de red neuronal artificial (ANN), focalizando así el modelo en las características que más influyen en las emisiones de CO_2 y mejorando su capacidad predictiva.

Esta correlación se puede evidenciar en la Figura 9 donde se especifica que las variables: tamaño del motor, número de cilindros, consumo de combustible en ciudad y consumo de combustible en avenidas son las que más influyen en las emisiones de CO_2 .

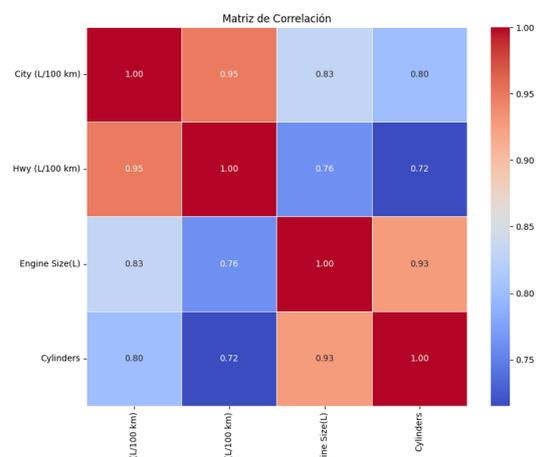


Figura 9. Correlación de variables que influyen en las emisiones de CO_2

Con la información derivada de la correlación, se avanza en la creación del modelo de aprendizaje automático utilizando una red neuronal artificial (ANN). En este proceso, se configura el modelo mediante la librería TensorFlow de Python, sobre la base de la siguiente estructura (Figura 10).

Donde, la estructura del modelo se define como secuencial, lo que implica una arquitectura de red neuronal donde las capas se disponen secuencialmente. Se

observan cuatro capas densamente conectadas (Dense), numeradas como dense, $dense_1$, $dense_2$, y $dense_3$. La primera capa tiene 64 neuronas, la segunda tiene 32, la tercera tiene 16, y la capa de salida tiene 1 neurona. Cada capa utiliza la función de activación ‘relu’ (Rectified Linear Unit), excepto la capa de salida que emplea la activación ‘linear’ que corresponde a las emisiones de CO_2 . Los parámetros se calculan automáticamente y se detallan en la Figura 10, donde se muestra la cantidad de parámetros entrenables (Trainable params) y la suma total de parámetros, que asciende a 9089. Estos parámetros representan los pesos y sesgos de la red neuronal que se ajustarán durante el proceso de entrenamiento para optimizar el rendimiento del modelo.

Model: "sequential"

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 64)	6464
dense_1 (Dense)	(None, 32)	2080
dense_2 (Dense)	(None, 16)	528
dense_3 (Dense)	(None, 1)	17
Total params: 9,089		
Trainable params: 9,089		
Non-trainable params: 0		

Figura 10. Estructura de la ANN

Posterior al entrenamiento del modelo, se presenta la Figura 11 que corresponde la pérdida durante el entrenamiento o model loss. Que especifica que cuantas más épocas pasan, se espera que la pérdida disminuya y la precisión aumente, lo que indica que el modelo está aprendiendo y mejorando en su capacidad para hacer predicciones precisas. Por lo tanto, la Figura 11 da a conocer que el modelo está correctamente entrenado.



Figura 11. Pérdida durante el entrenamiento

Para evaluar el desempeño de un modelo de aprendizaje automático en sus predicciones, es fundamental medir su precisión. Esto se logra mediante el uso de

métricas de rendimiento como la raíz cuadrada del error cuadrático medio (RMSE), el error cuadrático medio (MSE) y el coeficiente de determinación (R^2 cuadrado). Estas métricas ayudan a confirmar la precisión de los modelos de regresión [19], desempeñan un papel crucial en la evaluación y perfeccionamiento de los modelos de aprendizaje, permitiendo una comprensión más profunda de su capacidad para explicar y predecir datos, estas métricas se presentan a continuación en la ecuación (2):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{real}^{(i)} - y_{pred}^{(i)})^2$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{real}^{(i)} - y_{pred}^{(i)})^2}$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Donde SS_{res} se define como la suma de los cuadrados de las discrepancias entre los valores observados y los valores predichos. Por otro lado, SS_{tot} representa la suma de los cuadrados de las diferencias entre los valores observados y la media de dichos valores. Un valor de R^2 cercano a 1 indica un buen ajuste del modelo, mientras que un valor cercano a 0 indica que el modelo no explica la variabilidad de los datos.

Los resultados de las métricas del modelo son altamente prometedores: el coeficiente de determinación (R^2) alcanza un destacado 0,992, indicando una capacidad sobresaliente para explicar la variabilidad en los datos. El error cuadrático medio (MSE) se sitúa en 20,59, lo cual demuestra una magnitud promedio de errores cuadráticos razonablemente baja, mientras que la raíz del error cuadrático medio (RMSE) de 4,53 confirma una precisión significativa en las predicciones del modelo. Estos resultados sugieren una fuerte capacidad del modelo para prever las emisiones de gases contaminantes en el parque automotor, respaldando su idoneidad para aplicaciones predictivas, estos valores se los evidencia en la Figura 12.

Coficiente de Determinación (R^2): 0.9922332006414919
 Error Cuadrático Medio (MSE): 20.59925370801521
 Raíz del Error Cuadrático Medio (RMSE): 4.538640072534416

Figura 12. Resultados de las métricas del modelo

Para evaluar la eficacia del modelo desarrollado, se llevó a cabo una prueba que comparó los valores reales de las emisiones de gases contaminantes con las predicciones generadas por la red neuronal (Figura 13). En este análisis visual se buscó la proximidad de los puntos a una línea diagonal ideal, la cual representaría una predicción perfecta. La dispersión y la distribución de los puntos en la gráfica permiten una evaluación rápida de la capacidad del modelo para capturar la variabilidad en los datos reales. Una alineación cercana a la diagonal indica una predicción precisa, mientras que una dispersión significativa podría indicar áreas de mejora en la capacidad predictiva del modelo. Este

enfoque proporciona una evaluación intuitiva y visual de la calidad de las predicciones del modelo en relación con los datos reales.

Por lo tanto, la implementación de una ANN en el monitoreo de emisiones de gases contaminantes en vehículos presenta mejoras sustanciales y complementa eficazmente los métodos de monitoreo convencionales. A diferencia de enfoques más simples y lineales, la ANN pueden capturar relaciones no lineales complejas en los datos, ofreciendo una precisión mejorada en las predicciones. Su capacidad para adaptarse dinámicamente a cambios proporciona un monitoreo más robusto y adaptable. Además, la ANN pueden manejar eficientemente datos multidimensionales y complejos, procesando simultáneamente múltiples entradas, como temperatura, humedad y diversas emisiones de gases. La capacidad de aprendizaje automático y mejora continua permite una evolución constante en la precisión a medida que se acumulan más datos sobre emisiones vehiculares.

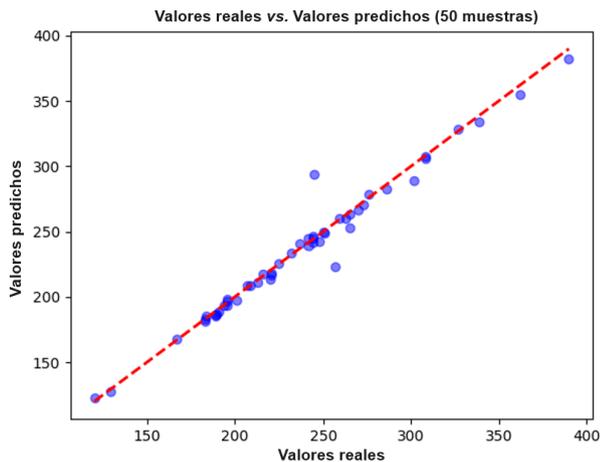


Figura 13. Prueba del modelo

4. Conclusiones

El presente artículo aborda de manera integral el monitoreo de emisiones de gases contaminantes en vehículos, desde la construcción del nodo sensor dedicado a la toma de datos hasta la creación de un modelo de red neuronal artificial (ANN) para predecir las emisiones de CO_2 . La implementación del nodo sensor, equipado con sensores DHT22, MQ9 y MQ135, demostró su eficacia al capturar datos clave como temperatura, humedad y concentraciones de gases durante las pruebas en un vehículo SUV Kia Sportage 2012. La validación de estos datos, realizada a través de una revisión técnica vehicular, respalda la precisión y confiabilidad de las mediciones.

Posterior, se aplicó un modelo de ANN utilizando datos de emisiones vehiculares de Canadá, centrándose en variables claves identificadas mediante correlación.

Los resultados del modelo, con un coeficiente de determinación (R^2) del 99,2 %, indican una capacidad efectiva para prever las emisiones de CO_2 . Estos hallazgos sugieren que la combinación de tecnologías de sensores avanzados y modelos de aprendizaje automático puede ofrecer una solución efectiva para el monitoreo y predicción de emisiones vehiculares, contribuyendo así a la gestión y mitigación de la contaminación ambiental.

Como trabajo futuro la arquitectura propuesta se implementará en los buses urbanos, donde la ANN podría contribuir significativamente al control de la calidad del aire y estado vehicular, permitiendo una gestión más efectiva de las políticas de circulación y planificación.

Referencias

- [1] J. Krause, C. Thiel, D. Tsokolis, Z. Samaras, C. Rota, A. Ward, P. Prenninger, T. Coosemans, S. Neugebauer, and W. Verhoeve, “EU road vehicle energy consumption and CO_2 emissions by 2050 – Expert-based scenarios,” *Energy Policy*, vol. 138, p. 111224, 2020. [Online]. Available: <https://doi.org/10.1016/j.enpol.2019.111224>
- [2] M. M. Ajmal, M. Khan, M. K. Shad, H. AlKatheeri, and F. Jabeen, “Empirical examination of societal, financial and technology-related challenges amid COVID-19 in service supply chains: evidence from emerging market,” *The International Journal of Logistics Management*, vol. 34, no. 4, pp. 994–1019, Jan 2023. [Online]. Available: <https://doi.org/10.1108/IJLM-04-2021-0220>
- [3] J. Lynn and N. Peeva, “Communications in the IPCC’s Sixth Assessment Report cycle,” *Climatic Change*, vol. 169, no. 1, p. 18, Nov 2021. [Online]. Available: <https://doi.org/10.1007/s10584-021-03233-7>
- [4] Z. Yang and A. Bandivadekar, *Light-duty vehicle greenhouse gas and fuel economy standards*. The International Council on clean Transportation, 2017. [Online]. Available: <https://bit.ly/4anzh8u>
- [5] R. Guensler, “Data needs for evolving motor vehicle emission modeling approaches,” *The University of California, Transportation Center*, 1993. [Online]. Available: <https://bit.ly/3THURO5>
- [6] Y. Lu, *Traffic-Related PM_{2.5} Air Pollution and Schools in Proximity to Major Roadways in Shanghai, China*. University of Washington, Department of Urban Design and Planning, 2016. [Online]. Available: <https://bit.ly/43KjOgq>
- [7] N. Kozarev and N. Ilieva, “Plume rise in particular meteorological conditions,” *Journal*

- of the University of Chemical Technology and Metallurgy, vol. 46, pp. 305–308, 01 2011. [Online]. Available: <https://bit.ly/3VLqMQ3>
- [8] N. Barmpareos, V. D. Assimakopoulos, M. N. Assimakopoulos, and E. Tsairidi, “Particulate matter levels and comfort conditions in the trains and platforms of the Athens underground metro,” *AIMS Environmental Science*, vol. 3, no. 2, pp. 199–219, 2016. [Online]. Available: <https://doi.org/10.3934/environsci.2016.2.199>
- [9] R. Senthilkumar, P. Venkatakrishnan, and N. Balaji, “Intelligent based novel embedded system based IoT enabled air pollution monitoring system,” *Microprocessors and Microsystems*, vol. 77, p. 103172, 2020. [Online]. Available: <https://doi.org/10.1016/j.micpro.2020.103172>
- [10] L. Moses, Tamilselvan, Raju, and Karthikeyan, “IoT enabled Environmental Air Pollution Monitoring and Rerouting system using Machine learning algorithms,” *IOP Conference Series: Materials Science and Engineering*, vol. 955, no. 1, p. 012005, nov 2020. [Online]. Available: <https://dx.doi.org/10.1088/1757-899X/955/1/012005>
- [11] V. Behal and R. Singh, “Personalised healthcare model for monitoring and prediction of airpollution: machine learning approach,” *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 33, no. 3, pp. 425–449, 2021. [Online]. Available: <https://doi.org/10.1080/0952813X.2020.1744197>
- [12] C. Shetty, B. Sowmya, S. Seema, and K. Srinivasa, “Chapter eight - air pollution control model using machine learning and IoT techniques,” in *The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases*, ser. Advances in Computers, P. Raj and P. Evangeline, Eds. Elsevier, 2020, vol. 117, no. 1, pp. 187–218. [Online]. Available: <https://doi.org/10.1016/bs.adcom.2019.10.006>
- [13] “Guest Editorial: Special Section on Integration of Big Data and Artificial Intelligence for Internet of Things,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2562–2565, 2020. [Online]. Available: <https://doi.org/10.1109/TII.2019.2958638>
- [14] R. Mumtaz, S. M. H. Zaidi, M. Z. Shakir, U. Shafi, M. M. Malik, A. Haque, S. Mumtaz, and S. A. R. Zaidi, “Internet of Things (IoT) Based Indoor Air Quality Sensing and Predictive Analytic—A COVID-19 Perspective,” *Electronics*, vol. 10, no. 2, 2021. [Online]. Available: <https://doi.org/10.3390/electronics10020184>
- [15] M. Abdel-Basset, G. Manogaran, M. Mohamed, and E. Rushdy, “Internet of things in smart education environment: Supportive framework in the decision-making process,” *Concurrency and Computation: Practice and Experience*, vol. 31, no. 10, p. e4515, 2019. [Online]. Available: <https://doi.org/10.1002/cpe.4515>
- [16] M. Montaña Blacio, J. González-Escarabay, O. Jiménez-Sarango, L. Mingo-Morocho, and C. Carrión-Aguirre, “Design and deployment of an IoT-based monitoring system for hydroponic crops,” *Ingenius. Revista de Ciencia y Tecnología*, no. 30, pp. 9–18, 2023. [Online]. Available: <https://doi.org/10.17163/ings.n30.2023.01>
- [17] M. Montaña Blacio, V. García Santos, D. Jaramillo Chamba, W. Torres Guin, and L. Chuquimarca Jiménez, “Empowering Low-Power Wide-Area Networks: Unlocking the Potential of Sigfox in Local Transmission,” in *Advanced Research in Technologies, Information, Innovation and Sustainability*, T. Guarda, F. Portela, and J. M. Diaz-Nafria, Eds. Cham: Springer Nature Switzerland, 2024, pp. 417–429. [Online]. Available: https://doi.org/10.1007/978-3-031-48930-3_32
- [18] J. K. Segura Gómez, *Prototipo de un sistema IoT para medición de gases de efecto invernadero*. Universidad Santo Tomás. Colombia, 2021. [Online]. Available: <https://bit.ly/3J4OvDl>
- [19] Kaggle. (2020) CO₂ Emission by Vehicles. Kaggle. [Online]. Available: <https://bit.ly/3J7Navw>