







Application of Machine Learning and Deep Learning in the analysis of air quality through IoT: a systematic review

Ocaña-Velásquez, Jesús Daniel¹, Dr. ; Castro-García, José Heiner², Mg. ; Miranda-Saldaña, Rodolfo Junior³, Mg. 
^{1,3}Universidad Tecnológica del Perú, Chimbote, Perú, c25777@utp.edu.pe, c20312@utp.edu.pe, e20206@utp.edu.pe

Abstract– The growing concern about climate change and environmental deterioration has made air pollution a global problem. This article presents a systematic review on the application of Machine Learning and Deep Learning techniques in air quality assessment. The objective of this research is to evaluate the effectiveness of Machine Learning and Deep Learning techniques in air quality analysis, identifying those that show superior performance in terms of accuracy and reliability of results. The PRISMA method was applied to compile 65 relevant articles on air quality. The findings indicate that Machine Learning and Deep Learning are crucial in this area, especially in research from India and China. The most common methods in Machine Learning are SVM and Random Forest, while in Deep Learning LSTM and CNN stand out. It is concluded that Machine Learning and Deep Learning are essential to assess air quality using IoT, with Machine Learning standing out for its accessibility and ease of interpretation in small data sets, while Deep Learning, despite requiring more resources and data, provides greater accuracy in the analysis.

Keywords– Machine Learning, Deep Learning, Air quality, Internet of Things (IoT).

Aplicación de Machine Learning y Deep Learning en el análisis de la calidad del aire a través de IoT: una revisión sistemática

Ocaña-Velásquez, Jesús Daniel¹, Dr. ; Castro-García, José Heiner², Mg. ; Miranda-Saldaña, Rodolfo Junior³, Mg. 
^{1,3}Universidad Tecnológica del Perú, Chimbote, Perú, c25777@utp.edu.pe, c20312@utp.edu.pe, e20206@utp.edu.pe

Resumen— *La creciente preocupación por el cambio climático y el deterioro ambiental ha hecho de la contaminación del aire un problema global. Este artículo presenta una revisión sistemática sobre la aplicación de técnicas de Machine Learning y Deep Learning en la evaluación de la calidad del aire. El objetivo de esta investigación es evaluar la eficacia de las técnicas de Machine Learning y Deep Learning en el análisis de la calidad del aire, identificando aquellas que muestran un rendimiento superior en términos de precisión y fiabilidad de los resultados. Se aplicó el método PRISMA para compilar 65 artículos relevantes sobre la calidad del aire. Los hallazgos indican que Machine Learning y Deep Learning son cruciales en esta área, especialmente en investigaciones de la India y China. Los métodos más comunes en Machine Learning son SVM y Bosque aleatorio, mientras que en Deep Learning destacan LSTM y CNN. Se concluye que Machine Learning y Deep Learning son fundamentales para evaluar la calidad del aire mediante IoT, destacándose Machine Learning por su accesibilidad y facilidad de interpretación en conjuntos de datos pequeños, mientras que Deep Learning, a pesar de requerir más recursos y datos, proporciona mayor precisión en el análisis.*

Palabras clave— *Machine Learning, Deep Learning, Calidad del aire, Internet de las Cosas (IoT).*

I. INTRODUCCIÓN

La tecnología basada en la Inteligencia Artificial está experimentando una transformación continua, especialmente en lo que respecta a las técnicas de Machine Learning y Deep Learning [1], [2]. Estas técnicas son fundamentales para analizar y predecir los problemas relacionados con la contaminación del aire, la cual aumenta significativamente el riesgo de desarrollar trastornos respiratorios, cáncer de pulmón y accidentes cerebrovasculares [3], [4], [5].

La contaminación ambiental es el resultado de la presencia de diversos compuestos químicos y biológicos en el aire de las principales ciudades del país [6], [7]. Con el aumento constante del número de vehículos en las áreas urbanas, las emisiones tóxicas se intensifican, deteriorando la calidad del aire en la región [8], [9].

El desarrollo de tecnologías IoT, junto con el análisis de datos mediante Machine Learning y Deep Learning, ofrece una oportunidad única para monitorear en tiempo real la calidad del aire. Esto permite detectar, predecir y gestionar de manera más eficiente la contaminación, mejorando la salud pública y la sostenibilidad ambiental [10], [11].

Los métodos de Machine Learning son fundamentales para el desarrollo de esta disciplina. Entre los más utilizados se encuentran: la regresión lineal, Bayesiano ingenuo, las máquinas de soporte vectorial (SVM), las redes neuronales, los árboles de decisión, el bosque aleatorio y el método de K vecinos más cercanos (KNN). Cada uno de estos métodos ofrece características únicas que los hacen adecuados para diferentes tipos de problemas y conjuntos de datos [12].

El Deep Learning es una subdisciplina del Machine Learning que emplea redes neuronales con múltiples capas. Incluye las Redes Neuronales Convolucionales (CNN) para el procesamiento de imágenes, las Redes Neuronales Recurrentes (RNN) para el análisis de secuencias y las Redes Neuronales de Memoria a Largo Plazo (LSTM), que optimizan la capacidad de las RNN para recordar información y gestionar dependencias a largo plazo [13], [14].

La preocupación por la calidad del aire y su impacto en la salud ha promovido el uso de sensores conectados a Internet para el monitoreo en tiempo real de la contaminación. Sin embargo, la gran cantidad de datos generados necesita técnicas de Machine Learning y Deep Learning para su análisis. A pesar de los avances, se carece de un marco sistemático que evalúe integralmente la efectividad de estas tecnologías, ya que las investigaciones se enfocan en aplicaciones específicas y no ofrecen una visión general sobre su implementación y eficacia.

El objetivo de esta investigación es evaluar la efectividad de las técnicas de Machine Learning y Deep Learning en el análisis de la calidad del aire, mediante la identificación y comparación de los métodos que destacan por su precisión y fiabilidad. A través de un enfoque sistemático, esta investigación pretende contribuir al desarrollo de herramientas y modelos más eficientes para la monitorización y gestión de la calidad del aire.

Este trabajo se divide en cinco secciones: la primera revisa estudios previos, la segunda describe la metodología utilizada, la tercera presenta los resultados, la cuarta discute estos hallazgos y finalmente se detallan las conclusiones.

II. TRABAJO RELACIONADO

El estudio [15] utiliza el método SVM para clasificar el estado de los sensores como normales o anormales. Si se detecta una anomalía, se activa un timbre para alertar sobre

posibles problemas. Los investigadores concluyeron que este sistema se integraría en la capa de middleware del IoT, permitiendo a las organizaciones anticipar la contaminación del aire y tomar decisiones informadas para mitigar su impacto en la población.

La investigación [16] ha diseñado un algoritmo para la extracción de datos que almacena información en una base de datos en la nube. Este trabajo de investigación desarrolló un terminal electrónico con un circuito Wi-Fi, utilizando tecnologías de Internet de las Cosas (IoT). Se implementó el método KNN, útil para analizar la calidad del aire y proporcionar datos relevantes para políticas ambientales y estudios de salud pública. Concluyó que el diseño permite detectar datos sobre contaminación y recopilar información meteorológica, que se publica en la base de datos en la nube.

De acuerdo al estudio [17] realizó un análisis sistemático de los modelos de Deep Learning aplicados a la predicción de la calidad del aire. Se evaluaron las metodologías tradicionales en comparación con los avances más recientes en Deep Learning. Los hallazgos revelaron que el Deep Learning ha adquirido una importancia significativa, debido a su habilidad para identificar relaciones no lineales complejas en grandes volúmenes de datos.

El estudio [18] desarrolló un sistema automatizado basado en IoT para clasificar la calidad del aire, evaluando el rendimiento de siete algoritmos de aprendizaje automático en dos escenarios: uno sin valores faltantes y otro con imputación mediante KNN. Se concluyó que optimizar el modelo con técnicas de aprendizaje automático y profundo en conjuntos apilados es esencial para obtener resultados más robustos y generalizables, especialmente al tratar con datos de alta dimensionalidad.

III. METODOLOGÍA

En esta investigación, se emplea el método PRISMA, el cual facilita la documentación clara y concisa de la información proveniente de artículos relevantes sobre este tema.

A. Preguntas de Investigación

Se realizaron cuatro preguntas de investigación utilizando el modelo PICO.

- RQ1: ¿Cuáles son las principales diferencias entre las características de Machine Learning y Deep Learning en el análisis de la calidad del aire?
- RQ2: ¿Cuáles son los métodos más utilizados en Machine Learning y Deep Learning para el análisis de la calidad del aire?
- RQ3: ¿Cuáles son las principales limitaciones de las técnicas de Machine Learning y Deep Learning en la evaluación de la calidad del aire?
- RQ4: ¿Como la integración de técnicas de Machine Learning y Deep Learning contribuye a mejorar el análisis de la calidad del aire?

B. Estrategia de búsqueda

La investigación se realiza mediante una estrategia de búsqueda detallada y el uso de filtros en bases de datos de prestigio, como Scopus, ScienceDirect, Web of Science, IEEE Xplore y EBSCOhost, con el propósito de identificar artículos pertinentes. A continuación, se aplican criterios de inclusión y exclusión, conforme a lo establecido en la Tabla 1 de acuerdo con la declaración PRISMA, para determinar cuáles artículos serán seleccionados para su posterior análisis en este estudio.

La estrategia de búsqueda utilizada para identificar artículos que incluyan tanto el título como las palabras clave del estudio es la siguiente:

("air quality" OR "global warming" OR "environmental pollution" OR "air pollution" OR "climate change" OR environment) AND ("artificial intelligence" OR "machine learning" OR "Deep Learning") AND (IoT OR " Internet of Things") AND (forecast OR predict OR analyze).

C. Criterios de inclusión y exclusión

Los criterios de inclusión y exclusión utilizados en esta revisión sistemática se describen de forma detallada en la Tabla 1.

TABLA I
CRITERIOS DE INCLUSIÓN Y EXCLUSIÓN

Criterios		
Inclusión	I01	Artículos relacionados con Maching Learning o Deep Learning en el análisis de la calidad de aire
	I02	Artículos que abordan el análisis de la calidad del aire mediante la aplicación de Machine Learning o Deep Learning a través de IoT.
	I03	Artículos en inglés
	I04	Artículos de los últimos 4 años (2021 -2024).
Exclusión	E01	Artículos que no guardan relación con el tema de investigación
	E02	Artículos que no tengan Open Access
	E03	Artículos que no están relacionados con Machine Learning o Deep Learning aplicados a la calidad del aire

Se utilizó el método PRISMA en tres etapas para seleccionar artículos de cinco bases de datos, con un total de 3110 artículos. Tras eliminar 242 duplicados quedando 2868, se redujo a 2015 al seleccionar artículos de acceso abierto. Se revisaron exhaustivamente, descartando 507 que no cumplieran criterios de inclusión, lo que dejó 346 para evaluación. De estos, se eliminaron 281 por falta de relación con las preguntas de investigación, resultando en 65 artículos seleccionados para el desarrollo de la investigación.

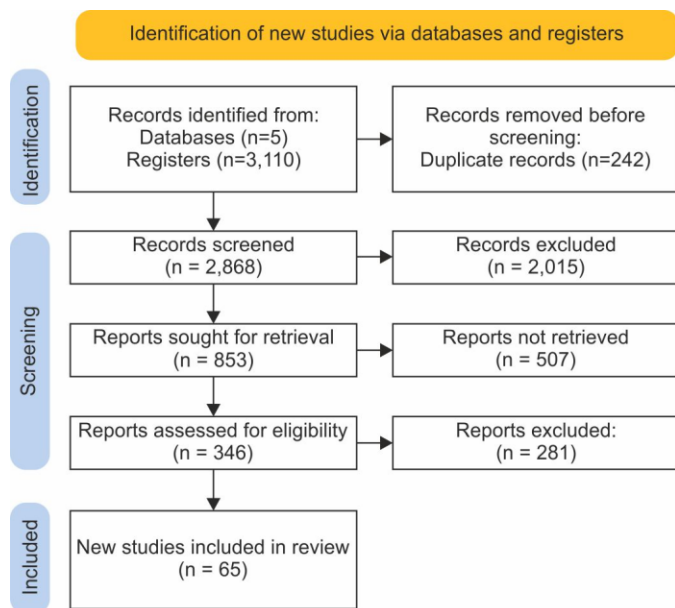


Fig. 1 Selección de artículos científicos según método PRISMA

IV. RESULTADOS

Se lleva a cabo un análisis bibliométrico de 65 artículos que satisfacen los criterios de inclusión establecidos.

A. Análisis bibliométrico

El análisis bibliométrico es una técnica cuantitativa que utiliza herramientas matemáticas y estadísticas para identificar tendencias en un campo de estudio. Mediante VOSViewer, se realizó un análisis de investigaciones, creando redes de palabras clave y mapas visuales que facilitan la interpretación de los datos. Se examinaron 235 palabras clave, lo que llevó a la identificación de 19 terminologías clave

Al analizar el mapa de la red que se muestra en la Figura 2, se identifican 19 palabras clave que presentan interconexiones significativas. Por ejemplo, el nodo "Air Quality" se relaciona estrechamente con "Machine Learning", "Deep Learning", "air pollution", "algorithm", "air monitoring" y "prediction". Por su parte, "Machine Learning" presenta conexiones con "Air Quality", "Deep Learning", "internet of things", "sensors", "artificial intelligence", "artificial neural", "air quality monitoring", "air pollution" y "prediction". Asimismo, el nodo "Deep Learning" se relaciona con "Air Quality", "Machine Learning", "artificial intelligence", "artificial neural", "sensors", "air pollution prediction", "air pollution" y "sensor data". Finalmente, el nodo "internet of things" está conectado con "Machine Learning", "Deep Learning", "artificial intelligence", "artificial neural" y "air quality prediction".

El nodo "Air Quality" está en el centro de una red conectándose con "Machine Learning", "Deep Learning" e "Internet of Things", lo que destaca la importancia de utilizar estas tecnologías para analizar la calidad del aire.

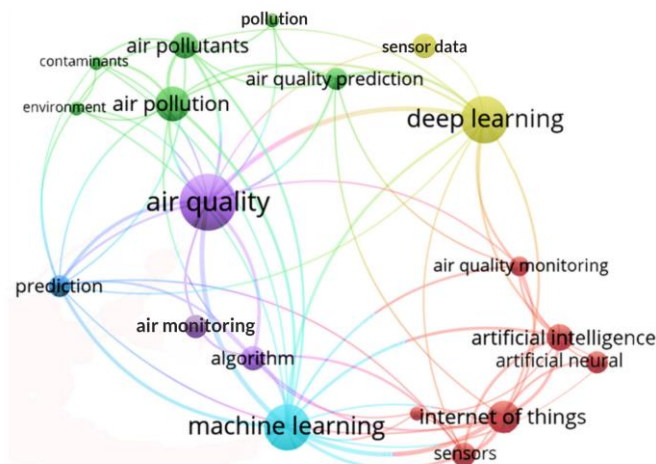


Fig. 2 Mapa bibliométrico de las relaciones entre palabras clave

B. Análisis de manuscritos

Se realizó una revisión de artículos en cinco bases de datos: Scopus, ScienceDirect, Web of Science, EBSCO Host e IEEE Xplore. En un inicio, se localizaron 3,110 estudios, de los cuales se eliminaron 242 duplicados. Tras aplicar los criterios de inclusión y exclusión, se seleccionaron 65 estudios, como se muestra en la Figura 3.

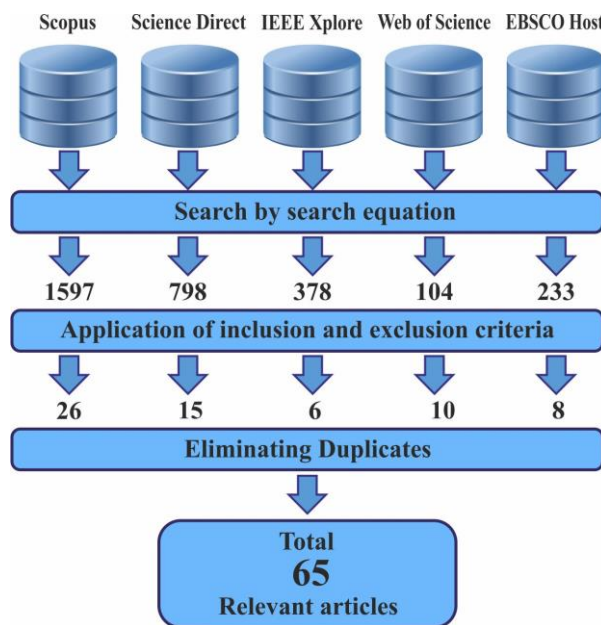


Fig. 3 Resultados obtenidos en la búsqueda.

La Figura 4 presenta una representación detallada del porcentaje de contribución informativa de diversas bases de datos, destacando Scopus, que lidera con un notable 40% de la contribución total. Seguidamente, Science Direct se posiciona con un 23%. Por su parte, Web of Science aporta un 16%. EBSCO Host y IEEE Xplore, aunque con contribuciones

menores, con un 12% y un 9%, respectivamente. Esta distribución porcentual no solo resalta la influencia de cada base de datos, sino que también refleja las preferencias y tendencias actuales en el acceso a la información científica.

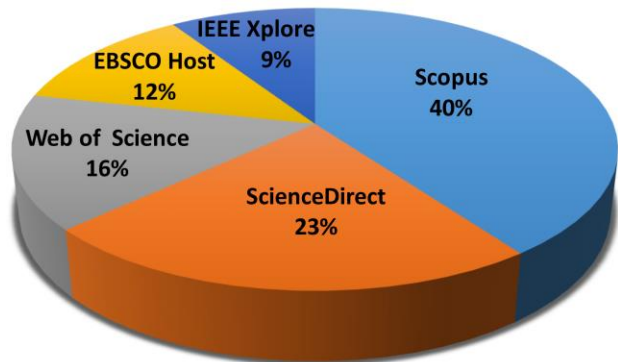


Fig. 4 Gráfico de resultados obtenidos en la búsqueda

Este estudio se centra en la cuantificación anual de artículos publicados, seleccionados según criterios de inclusión, desde 2021 hasta 2024. La Figura 5 muestra un gráfico de barras que representa la cantidad y el porcentaje de artículos incluidos en la revisión sistemática, clasificados según su año de publicación. Este crecimiento indica un aumento notable en la producción literaria sobre el tema, lo que muestra una mayor atención y relevancia en esta área de investigación en los últimos años. Se observa un significativo aumento en el número de publicaciones durante este período, alcanzando su punto máximo en 2024 con 26 artículos, lo que equivale al 37.14% del total.

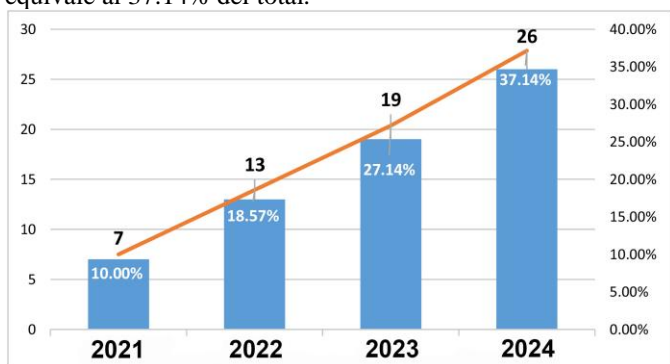


Fig. 5 Porcentajes de cantidad de artículos por año

La Figura 6 ilustra el número de manuscritos publicados anualmente en cada una de las bases de datos analizadas. En 2024, se evidenció un aumento notable en la cantidad de publicaciones en las cinco bases de datos. Scopus se posicionó a la cabeza con un total de 10 manuscritos, seguido de ScienceDirect con 6 y Web of Science con 4. Por su parte, tanto IEEE Xplore como EBSCO Host registraron 3 publicaciones cada uno.

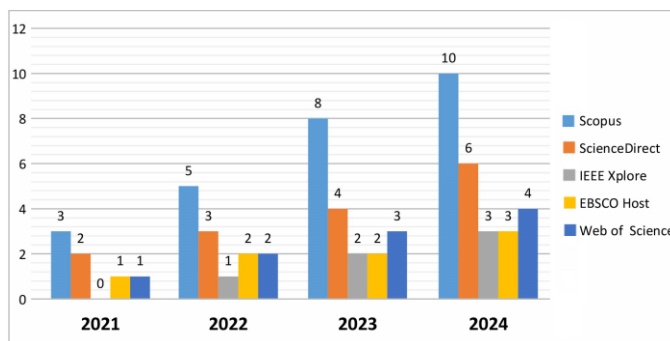


Fig. 6 Artículos por año y base de datos

La Figura 7 ilustra la distribución de artículos publicados por país, destacando a India como el líder en contribución a la investigación, con un total de 25 artículos que reflejan su creciente influencia en el ámbito académico. A continuación, China ocupa el segundo lugar con 15 artículos, lo que evidencia su compromiso con la producción de conocimiento científico, seguido de cerca por Corea del Sur, que ha aportado 5 artículos, mostrando también su dedicación a la investigación. El Reino Unido, con una contribución de 4 artículos, y España, con 3. Estados Unidos, Arabia Saudita y Egipto, aunque con una producción más modesta, han publicado 2 artículos cada uno. Por último, los países con la menor cantidad de publicaciones son Islandia, Italia, Turquía, Irak, Australia, Grecia y Japón, cada uno con un solo artículo publicado.

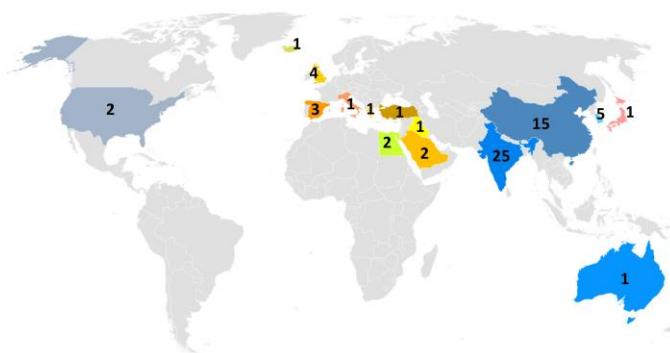


Fig. 7 Número de artículos publicados por país

V. DISCUSIÓN

Se realizó un riguroso proceso de selección y revisión de literatura con el fin de determinar cuántos estudios han incorporado técnicas de Machine Learning o Deep Learning en el análisis de la calidad del aire. Este meticuloso enfoque no solo facilitó la identificación de tendencias y vacíos en la investigación existente, sino que también permitió abordar de manera efectiva las preguntas de investigación formuladas en nuestro estudio. Al profundizar en esta revisión, se busca contribuir a un entendimiento más amplio sobre la aplicación

de estas tecnologías avanzadas en el monitoreo y mejora de la calidad del aire.

RQ1: ¿Cuáles son las principales diferencias entre las características de Machine Learning y Deep Learning en el análisis de la calidad del aire?

La Tabla 2 presenta un resumen conciso de las diferencias fundamentales entre Machine Learning y Deep Learning, destacando que Machine Learning requiere menos datos [19], lo que favorece un entrenamiento más ágil y una interpretación más accesible. Por otro lado, Deep Learning demanda grandes volúmenes de datos [20], lo que conlleva un mayor consumo de recursos y una complejidad en su interpretación. Esta distinción no solo subraya las ventajas de Machine Learning en entornos donde los datos son escasos, sino que también pone de relieve el potencial del Deep Learning para desarrollar modelos más sofisticados en contextos con abundancia de datos. Así, la elección entre Machine Learning y Deep Learning debe basarse en la disponibilidad de datos y en los objetivos específicos del análisis.

TABLA II
DIFERENCIAS ENTRE MACHINE LEARNING Y DEEP LEARNING

Campos de la IA	Características	Referencia
Machine Learning	Requiere menos Datos	[17], [16], [21], [22], [15], [23], [24], [25], [26], [19], [27], [28], [29], [30], [31], [32], [33], [34].
	Menos recursos de Pc	[17], [23], [19], [27], [32], [35].
	Entrenamiento más Rápida	[36], [17], [21], [23], [19], [27], [30], [31], [32], [33], [37], [38].
	Fáciles de interpretar	[15], [23], [24], [25], [26], [28], [31], [37], [38], [39], [40].
Deep Learning	Grandes volúmenes de datos	[30], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59].
	Utiliza más recursos computacionales	[17], [44], [45], [47], [48], [50], [51], [52], [56], [60], [61].
	Mayor Entrenamiento	[17], [42], [44], [45], [62], [20].
	Complejo de interpretar	[30], [41], [43], [44], [47], [48], [63].

RQ2: ¿Cuáles son los métodos más utilizados en Machine Learning y Deep Learning para analizar la calidad del aire?

Las tecnologías de aprendizaje automático y aprendizaje profundo proporcionan información valiosa para mejorar la toma de decisiones y las predicciones. La elección del método adecuado depende de los resultados deseados. La Tabla 3 analiza los métodos de Machine Learning y Deep Learning utilizados en la evaluación de la calidad del aire. En Machine Learning, el método más popular es Support Vector Machine

(SVM), seguido del Bosque Aleatorio, ambos efectivos para clasificación y predicción. En Deep Learning, el modelo más utilizado es Long Short-Term Memory (LSTM), que maneja secuencias temporales, seguido de las Redes Neuronales Convolucionales (CNN), son también relevantes por su capacidad en el análisis de datos espaciales. Este análisis destaca las tendencias actuales en el uso de técnicas avanzadas para mejorar la precisión y eficacia en la monitorización de la calidad del aire, lo que es esencial para el desarrollo de políticas ambientales más informadas y efectivas.

TABLA III
TIPO DE MÉTODOS DE MACHINE LEARNING Y DEEP LEARNING

Campos de la IA	Métodos	Referencia
Machine Learning	Regresión Lineal	[21], [27], [28], [32].
	Bayesiano ingenuo (NB)	[28], [31].
	Máquinas de Soporte Vectorial (SVM)	[17], [15], [25], [27], [28], [31], [34], [38], [64], [65], [66], [67], [68], [69], [70], [71].
	Árboles de decisión	[17], [22], [23], [28], [32], [38], [66], [72], [71].
	Bosque aleatorio (RF)	[17], [22], [23], [28], [31], [32], [34], [73], [68], [72], [69], [74], [70], [71].
	K-Vecino más cercano (KNN)	[17], [16], [22], [28], [31], [33], [66], [67], [72].
	DBSCAN	[36], [75].
	Red neuronal artificial (ANN)	[26], [30], [37], [65], [74]
Deep Learning	XGBoost	[17], [31], [46], [69]
	Redes neuronales convolucionales (CNN)	[17], [26], [35], [38], [44], [48], [51], [60], [76], [61], [71].
	DNN	[38], [69], [71].
	Gated Recurrent Unit (GRU)	[38], [46], [68], [74], [71].
	Memoria a corto plazo y largo plazo (LSTM)	[17], [25], [38], [41], [43], [44], [45], [46], [48], [49], [51], [52], [54], [55], [57], [58], [59], [60], [61], [62], [63], [77], [68], [72], [69], [74], [70], [71].
	BiLSTM	[46], [47], [56], [70]
	Redes neuronales recurrentes (RNN)	[35], [41], [46], [74], [70]
	híbrido CNN-LSTM	[40], [46], [47], [50], [61], [78], [68].
	híbrido LSTM -GRU	[55].
	híbrido HCNN-LSTM	[79].

RQ3: ¿Cuáles son las limitaciones de Machine Learning y Deep Learning en analizar la calidad del aire en sistemas IoT?

La Tabla 4 presenta un análisis de artículos que abordan las limitaciones del Machine Learning y el Deep Learning en la evaluación de la calidad del aire dentro de sistemas IoT.

Estos estudios subrayan que el rendimiento del Machine Learning está intrínsecamente ligado a la calidad de los datos [45], así como la presencia de ruido en los mismos. En contraste, las restricciones asociadas al Deep Learning se enfocan en su necesidad de grandes volúmenes de datos para un entrenamiento efectivo [62], así como en los elevados costos de los recursos.

TABLA IV
LIMITACIONES DE MACHINE LEARNING Y DEEP LEARNING

Campos de la IA	Limitaciones	Referencia
Machine Learning	Dependencia de la calidad de los datos.	[18], [16], [15], [23], [25], [19], [28], [37], [38], [39], [46], [73].
	Predicciones inexactas	[24], [27], [75].
	Capacidad de aprendizaje	[18], [31], [32], [33], [35].
	Datos con ruido	[18], [26], [32], [33], [37], [38], [40], [64], [75], [73].
	Bajo costo en la implementación	[18], [21], [26], [34], [37], [38].
Deep Learning	Requerimientos de grandes volúmenes de datos	[39], [42], [44], [45], [46], [47], [48], [49], [51], [52], [53], [54], [55], [56], [58], [59], [60], [62], [20], [77].
	Costos de recursos	[17], [29], [47], [48], [50], [51], [58], [20], [63], [76].
	Dependencia de la calidad de datos	[41], [44], [45], [54], [57], [20].
	Datos recopilados con perturbación o ruido	[46], [47], [51], [52], [56], [77].
	Alto costo en la implementación	[56], [58].
	Más Tiempo de entrenamiento	[30], [43], [44], [45], [46], [48], [49], [60].

RQ4: ¿Como la integración de técnicas de Machine Learning y Deep Learning contribuye a mejorar el análisis de la calidad del aire?

La Tabla 5 presenta una integración de Machine Learning y Deep Learning que han demostrado ser fundamentales para mejorar el análisis de la calidad del aire. Se realizaron diversas pruebas y se evaluaron múltiples métodos de estas tecnologías, y cada uno de estos enfoques mostró mejoras en eficiencia y precisión en comparación con otros métodos. Sin embargo, los métodos más eficientes y precisos son XGBoost y el Bosque Aleatorio. La combinación de estas técnicas puede ser un enfoque efectivo para el monitoreo y análisis de la calidad del aire.

Esta revisión sistemática analiza la calidad del aire mediante diversas técnicas de Machine Learning y Deep Learning, resaltando un gran potencial para abordar desafíos en este campo, lo que podría facilitar la apertura de nuevas oportunidades para investigaciones futuras. A diferencia de estudios anteriores que se centraron en enfoques limitados,

este trabajo ofrece un análisis comparativo que resalta la efectividad de métodos como XGBoost y el Bosque Aleatorio. Además, al resumir resultados de diferentes pruebas, establece un marco de referencia útil para investigadores y profesionales en la selección de métodos más eficientes para el monitoreo de la calidad del aire, superando limitaciones de estudios previos.

TABLA V
LA INTEGRACIÓN DE TÉCNICAS DE MACHINE LEARNING Y DEEP LEARNING

Métodos		Método más eficiente	Precisión %	Referencia
Machine Learning	Deep Learning			
Bosque aleatorio, Árboles de decisión, KNN, SVM, XGBoost	CNN, SV-CNN	SV-CNN	90.63	[18]
Árboles de decisión, Bosque aleatorio, SVM.	DNN, GRU, LSTM, CNN.	GRU	97.00	[38]
XGBoost.	LSTM, BiLSTM, RNN, GRU, Bi-GRU, CNN-BiLSTM	XGBoost	98.29	[46]
Bosque aleatorio, MLP, SVM.	GRU, LSTM, Deep-LSTM, DM-LSTM, CNN-LSTM.	CNN-LSTM	98.17	[68]
Bosque aleatorio, Árbol de decisiones, KNN.	LSTM	LSTM	98.30	[72]
Bosque aleatorio, SVM, XGBoost,	DNN, LSTM.	XGBoost,	86.20	[69]
Bosque aleatorio, ANN, SVR,	RNN, LSTM, GRU.	Bosque aleatorio	94.00	[74]
Bosque aleatorio, Gradient Boost (GB), SVM,	LSTM, BiLSTM, RNN.	Bosque aleatorio	71.80	[70]
Bosque aleatorio, SVR, ELM-SO, GBR, XGBoost,	LSTM	ELM-SO	92.80	[71]

VI. CONCLUSIONES

La revisión sobre el uso de Machine Learning y Deep Learning en el análisis de la calidad del aire a través de IoT, resalta diferencias significativas entre ambas técnicas. Machine Learning es efectivo con conjuntos de datos pequeños, facilitando su entrenamiento e interpretación, mientras que Deep Learning ofrece mayor precisión, pero necesita más datos y poder computacional. La elección de la técnica adecuada debe depender de las necesidades del proyecto y las características de los datos, lo que optimiza la gestión de la calidad del aire y ayuda en la implementación de soluciones sostenibles para el monitoreo ambiental.

Los estudios indican que el método SVM es el más usado en Machine Learning para problemas de clasificación, especialmente en la calidad del aire, seguido por el Bosque

aleatorio, que destaca por su versatilidad y fácil uso. En el ámbito del Deep Learning, las LSTM son las más utilizadas debido a su habilidad para aprender patrones a largo plazo, seguidas por las CNN, que se destacan por su eficacia en el procesamiento de datos complejos. Estos hallazgos subrayan la importancia de seleccionar el método adecuado según las características del problema a resolver, lo cual es crucial para optimizar los resultados en el monitoreo y análisis de la calidad del aire.

Machine Learning y Deep Learning son herramientas poderosas para el análisis de datos, pero cada una presenta sus propias limitaciones. El Machine Learning depende en gran medida de la calidad de los datos, así como de la cantidad de ruido presente en ellos, lo que puede afectar la precisión de los modelos. Por otro lado, el Deep Learning requiere volúmenes de datos significativamente grandes y suele implicar altos costos en términos de recursos computacionales. Por lo tanto, la elección entre estas técnicas debe basarse en las necesidades específicas del proyecto, la disponibilidad de datos y el presupuesto, con el fin de garantizar la efectividad del análisis de la calidad del aire.

La integración de técnicas de Machine Learning y Deep Learning es fundamental para optimizar el estudio de la calidad del aire. Los resultados obtenidos demuestran que, si bien existen múltiples métodos que pueden mejorar la eficiencia y precisión del análisis, XGBoost y el Bosque Aleatorio emergen como los más efectivos en este contexto. Esta evidencia sugiere que la combinación de estas metodologías no solo proporciona un enfoque robusto, sino que también se presenta como una solución innovadora y prometedora para abordar los desafíos del monitoreo ambiental.

REFERENCIAS

[1] K. Moulaei *et al.*, “Explainable artificial intelligence (XAI) for predicting the need for intubation in methanol-poisoned patients: a study comparing deep and machine learning models,” *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/S41598-024-66481-4.

[2] M. Sajid *et al.*, “Enhancing intrusion detection: a hybrid machine and deep learning approach,” *J. Cloud Comput.*, vol. 13, no. 1, Dec. 2024, doi: 10.1186/S13677-024-00685-X.

[3] N. Wright *et al.*, “Long-term ambient air pollution exposure and cardio-respiratory disease in China: findings from a prospective cohort study,” *Environ. Heal. A Glob. Access Sci. Source*, vol. 22, no. 1, Dec. 2023, doi: 10.1186/S12940-023-00978-9.

[4] P. D. Lavrič, M. Ivanovski, D. Goričanec, and D. Urbanc, “Particulate matter air pollution in the Republic of Slovenia and its national spatial emissions release,” *J. Renew. Energies*, vol. 26, no. 2, pp. 123–139, Dec. 2023, doi: 10.54966/JREEN.V26I2.1149.

[5] J. Yang *et al.*, “Application of machine learning to predict hospital visits for respiratory diseases using meteorological and air pollution factors in Linyi, China,” *Environ. Sci. Pollut. Res.*, vol. 30, no. 38, pp. 88431–88443, Aug. 2023, doi: 10.1007/S11356-023-28682-8.

[6] A. Kallel, M. Ksibi, H. Ben Dhia, and N. Khélifi, “Pollutant removal and the health effects of environmental pollution,” *Environ. Sci. Pollut. Res.*, vol. 27, no. 19, pp. 23375–23378, Jul. 2020, doi: 10.1007/S11356-020-08775-4.

[7] M. Shalaby and F. S. Alorifi, “Environmental pollution detection: A novel chirped spectral modulation algorithm for a more accurate monitoring of gas pollutants in the atmosphere,” *J. Eur. Opt. Soc. Publ.*, vol. 19, no. 1, 2023, doi: 10.1051/JEOS/2023005.

[8] H. di He and H. O. Gao, “Particulate matter exposure at a densely populated urban traffic intersection and crosswalk,” *Environ. Pollut.*, vol. 268, Jan. 2021, doi: 10.1016/J.ENVPOL.2020.115931.

[9] O. Mislyuk, E. Khomenko, O. Yehorova, and L. Zhytska, “ASSESSING RISK CAUSED BY ATMOSPHERIC AIR POLLUTION FROM MOTOR VEHICLES TO THE HEALTH OF POPULATION IN URBANIZED AREAS,” *Eastern-European J. Enterp. Technol.*, vol. 1, no. 10(121), pp. 19–26, 2023, doi: 10.15587/1729-4061.2023.274174.

[10] L. García, A.-J. Garcia-Sanchez, R. Asorey-Cacheda, J. Garcia-Haro, and C.-L. Zúñiga-Cañón, “Smart Air Quality Monitoring IoT-Based Infrastructure for Industrial Environments,” *Sensors*, vol. 22, no. 23, 2022, doi: 10.3390/S22239221.

[11] N. Tsalikidis *et al.*, “Urban Traffic Congestion Prediction: A Multi-Step Approach Utilizing Sensor Data and Weather Information,” *Smart Cities*, vol. 7, no. 1, pp. 233–253, Feb. 2024, doi: 10.3390/SMARTCITIES7010010.

[12] G. Mostafa, H. Mahmoud, T. Abd El-Hafeez, and M. E. ElAraby, “Feature reduction for hepatocellular carcinoma prediction using machine learning algorithms,” *J. Big Data*, vol. 11, no. 1, Dec. 2024, doi: 10.1186/S40537-024-00944-3.

[13] Y. Li, “Analysis of English Classroom Teaching Behavior Mode in Environmental Protection Field Based on Deep Learning,” *Int. J. Comput. Intell. Syst.*, vol. 17, no. 1, Dec. 2024, doi: 10.1007/S44196-024-00457-0.

[14] L. Abualigah, Y. Y. Al-Ajlouni, M. S. Daoud, M. Altalhi, and H. Migdady, “Fake news detection using recurrent neural network based on bidirectional LSTM and GloVe,” *Soc. Netw. Anal. Min.*, vol. 14, no. 1, Dec. 2024, doi: 10.1007/S13278-024-01198-W.

[15] T. Shanmugaraja, M. Sakthivel, R. Shaila Shree, P. Sooraj, and K. Vishnu, “Retraction: Analysis of air quality using IoT with machine learning prediction,” in *Journal of Physics: Conference Series*, R. A., S. M.G., and O. C.C., Eds., India: IOP Publishing Ltd, 2021. doi: 10.1088/1742-6596/1916/1/012188.

[16] K. S. Krishna, T. Satish, and J. Mishra, “Machine Learning-Based IOT Air Quality and Pollution Detection,” *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, pp. 62 – 76, 2023, doi: 10.17762/ijrctec.v11i2s.6036.

[17] A. Aggarwal and D. Toshniwal, “A hybrid deep learning framework for urban air quality forecasting,” *J. Clean. Prod.*, vol. 329, p. 129660, 2021, doi: <https://doi.org/10.1016/j.jclepro.2021.129660>.

[18] K. Alnowaiser, A. A. Alarfaj, E. A. Alabdulqader, M. Umer, L. Cascone, and B. Alankar, “IoT based smart framework to predict air quality in congested traffic areas using SV-CNN ensemble and KNN imputation model,” *Comput. Electr. Eng.*, vol. 118, 2024, doi: 10.1016/j.compeleceng.2024.109311.

[19] R. Zayed and M. Abbod, “Breathable Cities: Dynamic Machine Learning Modelling Approaches for Advanced Air Pollution Control,” *Appl. Sci.*, vol. 14, no. 13, p. 5581, Jun. 2024, doi: 10.3390/app14135581.

[20] Z. Zhang, S. Zhang, C. Chen, and J. Yuan, “A systematic survey of air quality prediction based on deep learning,” *Alexandria Eng. J.*, vol. 93, pp. 128–141, 2024, doi: <https://doi.org/10.1016/j.aej.2024.03.031>.

[21] T. S. Anggraini, H. Irie, A. D. Sakti, and K. Wikantika, “Machine learning-based global air quality index development using remote sensing and ground-based stations,” *Environ. Adv.*, vol. 15, 2024, doi: 10.1016/j.envadv.2023.100456.

[22] I. A. Sulaimon, H. Alaka, R. Olu-Ajayi, M. Ahmad, S. Ajayi, and A. Hye, “Effect of traffic data set on various machine-learning algorithms when forecasting air quality,” *J. Eng. Des. Technol.*, vol. 22, no. 3, pp. 1030–1056, 2024, doi: 10.1108/JEDT-10-2021-0554.

[23] F. Tuluri, R. Remata, W. L. Walters, and P. B. Tchounwou, “Impact of Regional Mobility on Air Quality during COVID-19 Lockdown in Mississippi, USA Using Machine Learning,” *Int. J. Environ. Res. Public Health*, vol. 20, no. 11, 2023, doi: 10.3390/ijerph20116022.

[24] B. Zhang, M. Duan, Y. Sun, Y. Lyu, Y. Hou, and T. Tan, “Air Quality Index Prediction in Six Major Chinese Urban Agglomerations: A Comparative Study of Single Machine Learning Model, Ensemble Model, and Hybrid Model,” *Atmosphere (Basel)*, vol. 14, no. 10, 2023,

doi: 10.3390/atmos14101478.

- [25] N. N. Maltare and S. Vahora, "Air Quality Index prediction using machine learning for Ahmedabad city," *Digit. Chem. Eng.*, vol. 7, p. 100093, 2023, doi: <https://doi.org/10.1016/j.dche.2023.100093>.
- [26] S. Koziel, A. Pietrenko-Dabrowska, M. Wojcikowski, and B. Pankiewicz, "Efficient calibration of cost-efficient particulate matter sensors using machine learning and time-series alignment," *Knowledge-Based Syst.*, vol. 295, 2024, doi: [10.1016/j.knsys.2024.111879](https://doi.org/10.1016/j.knsys.2024.111879).
- [27] V. Devasekhar and P. Natarajan, "Prediction of Air Quality and Pollution using Statistical Methods and Machine Learning Techniques," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 4, pp. 927–937, 2023, doi: [10.14569/IJACSA.2023.01404103](https://doi.org/10.14569/IJACSA.2023.01404103).
- [28] A. Pant, S. Sharma, and K. Pant, "Evaluation of Machine Learning Algorithms for Air Quality Index (AQI) Prediction," *J. Reliab. Stat. Stud.*, vol. 16, no. 2, pp. 229 – 242, 2023, doi: [10.13052/jrss0974-8024.1621](https://doi.org/10.13052/jrss0974-8024.1621).
- [29] A. S. Moursi, N. El-Fishawy, S. Djahel, and M. A. Shouman, "An IoT enabled system for enhanced air quality monitoring and prediction on the edge," *Complex Intell. Syst.*, vol. 7, no. 6, pp. 2923–2947, 2021, doi: [10.1007/s40747-021-00476-w](https://doi.org/10.1007/s40747-021-00476-w).
- [30] S. Shin, K. Baek, and H. So, "Rapid monitoring of indoor air quality for efficient HVAC systems using fully convolutional network deep learning model," *Build. Environ.*, vol. 234, p. 110191, 2023, doi: <https://doi.org/10.1016/j.buildenv.2023.110191>.
- [31] K. K. Meena, D. Bairwa, and A. Agarwal, "A machine learning approach for unraveling the influence of air quality awareness on travel behavior," *Decis. Anal. J.*, vol. 11, p. 100459, 2024, doi: <https://doi.org/10.1016/j.dajour.2024.100459>.
- [32] S. Al-Eidi, F. Amsaad, O. Darwish, Y. Tashtoush, A. Alqahtani, and N. Niveshitha, "Comparative Analysis Study for Air Quality Prediction in Smart Cities Using Regression Techniques," *IEEE Access*, vol. 11, pp. 115140–115149, 2023, doi: [10.1109/ACCESS.2023.3323447](https://doi.org/10.1109/ACCESS.2023.3323447).
- [33] V. N. Matthaos, L. D. Knibbs, L. J. Kramer, L. R. Crilley, and W. J. Bloss, "Predicting real-time within-vehicle air pollution exposure with mass-balance and machine learning approaches using on-road and air quality data," *Atmos. Environ.*, vol. 318, p. 120233, Feb. 2024, doi: [10.1016/j.atmosenv.2023.120233](https://doi.org/10.1016/j.atmosenv.2023.120233).
- [34] S. A. Siddiqui, N. Fatima, and A. Ahmad, "Smart Air Pollution Monitoring System with Smog Prediction Model using Machine Learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 8, pp. 401 – 409, 2021, doi: [10.14569/IJACSA.2021.0120846](https://doi.org/10.14569/IJACSA.2021.0120846).
- [35] J. Wang, L. Jin, X. Li, S. He, M. Huang, and H. Wang, "A Hybrid Air Quality Index Prediction Model Based on CNN and Attention Gate Unit," *IEEE Access*, vol. 10, pp. 113343–113354, 2022, doi: [10.1109/ACCESS.2022.3217242](https://doi.org/10.1109/ACCESS.2022.3217242).
- [36] D. Komarasamy, G. C. Nandhidha, S. Nandhinidevi, P. L. Nanthini, Mohanasaranya, and K. Kousalya, "Air Quality Prediction and Classification using Machine Learning," in *Proceedings - 7th International Conference on Computing Methodologies and Communication, ICCMC 2023*, 2023, pp. 187 – 191. doi: [10.1109/ICCMC56507.2023.10083760](https://doi.org/10.1109/ICCMC56507.2023.10083760).
- [37] K. Joseph and M. Lakshminath, "Air quality prediction on IoT real-time sensor using supervised machine learning," in *AIP Conference Proceedings*, S. J., Ed., India, 2024. doi: [10.1063/5.0216495](https://doi.org/10.1063/5.0216495).
- [38] H. Zhou, H. Luo, K. K. L. Lau, X. Qian, C. Ren, and P. Chau, "Predicting Emergency Department Utilization among Older Hong Kong Population in Hot Season: A Machine Learning Approach," *Inf.*, vol. 13, no. 9, Sep. 2022, doi: [10.3390/INFO13090410](https://doi.org/10.3390/INFO13090410).
- [39] M. Méndez, M. G. Merayo, and M. Núñez, "Machine learning algorithms to forecast air quality: a survey," *Artif. Intell. Rev.*, vol. 56, no. 9, pp. 10031–10066, 2023, doi: [10.1007/s10462-023-10424-4](https://doi.org/10.1007/s10462-023-10424-4).
- [40] M. A. Khan, H.-C. Kim, and H. Park, "Leveraging Machine Learning for Fault-Tolerant Air Pollutants Monitoring for a Smart City Design," *Electron.*, vol. 11, no. 19, 2022, doi: [10.3390/electronics11193122](https://doi.org/10.3390/electronics11193122).
- [41] B. Das, Ö. O. Dursun, and S. Toraman, "Prediction of air pollutants for air quality using deep learning methods in a metropolitan city," *Urban Clim.*, vol. 46, p. 101291, 2022, doi: <https://doi.org/10.1016/j.uclim.2022.101291>.
- [42] W. Wang, X. An, Q. Li, Y. Geng, H. Yu, and X. Zhou, "Optimization research on air quality numerical model forecasting effects based on deep learning methods," *Atmos. Res.*, vol. 271, p. 106082, 2022, doi: <https://doi.org/10.1016/j.atmosres.2022.106082>.
- [43] W. Fang, R. Zhu, and J. C.-W. Lin, "An air quality prediction model based on improved Vanilla LSTM with multichannel input and multiroute output," *Expert Syst. Appl.*, vol. 211, p. N.PAG, 2023, [Online]. Available: <https://search.ebscohost.com/login.aspx?direct=true&db=aph&app=AN=159798737&lang=es&site=ehost-live&auth=es&custid=ns256095>
- [44] I. I. Prado-Rujas, A. García-Dopico, E. Serrano, M. L. Córdoba, and M. S. Pérez, "A multivariable sensor-agnostic framework for spatio-temporal air quality forecasting based on Deep Learning," *Eng. Appl. Artif. Intell.*, vol. 127, Jan. 2024, doi: [10.1016/J.ENGAPPAI.2023.107271](https://doi.org/10.1016/J.ENGAPPAI.2023.107271).
- [45] Y. Feng, "Air Quality Prediction Model Using Deep Learning in Internet of Things Environmental Monitoring System," *Mob. Inf. Syst.*, vol. 2022, 2022, doi: [10.1155/2022/7221157](https://doi.org/10.1155/2022/7221157).
- [46] I. Ayus, N. Natarajan, and D. Gupta, "Comparison of machine learning and deep learning techniques for the prediction of air pollution: a case study from China," *Asian J. Atmos. Environ.*, vol. 17, no. 1, Dec. 2023, doi: [10.1007/S44273-023-00005-W](https://doi.org/10.1007/S44273-023-00005-W).
- [47] P. D. Bharathi, A. N. Velu, and B. S. Palaniappan, "Design and Enhancement of a Fog-Enabled Air Quality Monitoring and Prediction System: An Optimized Lightweight Deep Learning Model for a Smart Fog Environmental Gateway," *Sensors*, vol. 24, no. 15, Aug. 2024, doi: [10.3390/S24155069](https://doi.org/10.3390/S24155069).
- [48] E. I. Fernandez, A. J. J. Valera, and J. T. F. Breis, "Improving Air Quality Zoning Through Deep Learning and Hyperlocal Measurements," *IEEE Access*, vol. 12, pp. 38700–38716, 2024, doi: [10.1109/ACCESS.2024.3374208](https://doi.org/10.1109/ACCESS.2024.3374208).
- [49] Y. Zhu, S. A. Al-Ahmed, M. Z. Shakir, and J. I. Olszewska, "LSTM-Based IoT-Enabled CO₂ Steady-State Forecasting for Indoor Air Quality Monitoring," *Electron.*, vol. 12, no. 1, Jan. 2023, doi: [10.3390/ELECTRONICS12010107](https://doi.org/10.3390/ELECTRONICS12010107).
- [50] W. Cao, R. Zhang, and W. Cao, "Multi-Site Air Quality Index Forecasting Based on Spatiotemporal Distribution and PatchTSTEnhanced: Evidence from Hebei Province in China," *IEEE Access*, 2024, doi: [10.1109/ACCESS.2024.3460187](https://doi.org/10.1109/ACCESS.2024.3460187).
- [51] Z. Guo *et al.*, "Optimized air quality management based on air quality index prediction and air pollutants identification in representative cities in China," *Sci. Rep.*, vol. 14, no. 1, Dec. 2024, doi: [10.1038/S41598-024-68972-W](https://doi.org/10.1038/S41598-024-68972-W).
- [52] A. T. Nguyen, D. H. Pham, B. L. Oo, Y. Ahn, and B. T. H. Lim, "Predicting air quality index using attention hybrid deep learning and quantum-inspired particle swarm optimization," *J. Big Data*, vol. 11, no. 1, Dec. 2024, doi: [10.1186/S40537-024-00926-5](https://doi.org/10.1186/S40537-024-00926-5).
- [53] C.-Y. Lin, Y.-S. Chang, and S. Abimannan, "Ensemble multifeatured deep learning models for air quality forecasting," *Atmos. Pollut. Res.*, vol. 12, no. 5, p. 101045, 2021, doi: <https://doi.org/10.1016/j.apr.2021.03.008>.
- [54] R. Janarthanan, P. Partheeban, K. Somasundaram, and P. Navin Elamparithi, "A deep learning approach for prediction of air quality index in a metropolitan city," *Sustain. Cities Soc.*, vol. 67, p. 102720, 2021, doi: <https://doi.org/10.1016/j.scs.2021.102720>.
- [55] N. Sarkar, R. Gupta, P. K. Keserwani, and M. C. Govil, "Air Quality Index prediction using an effective hybrid deep learning model," *Environ. Pollut.*, vol. 315, p. 120404, 2022, doi: <https://doi.org/10.1016/j.envpol.2022.120404>.
- [56] P. D. Bharathi, V. A. Narayanan, and P. B. Sivakumar, "Fog computing enabled air quality monitoring and prediction leveraging deep learning in IoT," *J. Intell. & Fuzzy Syst.*, vol. 43, no. 5, pp. 5621–5642, 2022, [Online]. Available: <https://search.ebscohost.com/login.aspx?direct=true&db=egs&app=AN=159498632&lang=es&site=ehost-live&auth=es&custid=ns256095>
- [57] M. Marzouk and M. Atef, "Assessment of Indoor Air Quality in Academic Buildings Using IoT and Deep Learning," *Sustain.*, vol. 14, no. 12, 2022, doi: [10.3390/su14127015](https://doi.org/10.3390/su14127015).
- [58] A. Ramachandran, K. Gayathri, R. H. Aswathy, and S. Zulaikha Beevi,

- “Modeling of internet of things enabled sustainable environment air pollution monitoring system,” *Glob. Nest J.*, vol. 25, no. 4, pp. 172 – 179, 2023, doi: 10.30955/gnj.004707.
- [59] I.-I. Prado-Rujas, A. García-Dopico, E. Serrano, M. L. Córdoba, and M. S. Pérez, “A multivariable sensor-agnostic framework for spatio-temporal air quality forecasting based on Deep Learning,” *Eng. Appl. Artif. Intell.*, vol. 127, p. 107271, 2024, doi: <https://doi.org/10.1016/j.engappai.2023.107271>.
- [60] K. Hu, X. Guo, X. Gong, X. Wang, J. Liang, and D. Li, “Air quality prediction using spatio-temporal deep learning,” *Atmos. Pollut. Res.*, vol. 13, no. 10, p. 101543, 2022, doi: <https://doi.org/10.1016/j.apr.2022.101543>.
- [61] X. Wang, M. Wang, X. Liu, Y. Mao, Y. Chen, and S. Dai, “Surveillance-image-based outdoor air quality monitoring,” *Environ. Sci. Ecotechnology*, vol. 18, p. 100319, Mar. 2024, doi: 10.1016/J.ESE.2023.100319.
- [62] W. Du *et al.*, “Deciphering urban traffic impacts on air quality by deep learning and emission inventory,” *J. Environ. Sci.*, vol. 124, pp. 745–757, 2023, doi: <https://doi.org/10.1016/j.jes.2021.12.035>.
- [63] K. Chatterjee *et al.*, “Future Air Quality Prediction using Long Short-Term Memory based on Hyper Heuristic Multi-Chain Model,” *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3441109.
- [64] K. S. Rautela and M. K. Goyal, “Transforming air pollution management in India with AI and machine learning technologies,” *Sci. Rep.*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/S41598-024-71269-7.
- [65] S. J. Livingston, S. D. Kanmani, A. S. Ebenezer, D. Sam, and A. Joshi, “An ensemble method for air quality monitoring and control using machine learning,” *Meas. Sensors*, vol. 30, 2023, doi: 10.1016/j.measen.2023.100914.
- [66] A. H. Almaliki, A. Derdour, and E. Ali, “Air Quality Index (AQI) Prediction in Holy Makkah Based on Machine Learning Methods,” *Sustain.*, vol. 15, no. 17, 2023, doi: 10.3390/su151713168.
- [67] S. K. Natarajan, P. Shanmurthy, D. Arockiam, B. Balusamy, and S. Selvarajan, “Optimized machine learning model for air quality index prediction in major cities in India,” *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-54807-1.
- [68] M. Malhotra, S. Walia, C. C. Lin, I. K. Aulakh, and S. Agarwal, “A systematic scrutiny of artificial intelligence-based air pollution prediction techniques, challenges, and viable solutions,” *J. Big Data*, vol. 11, no. 1, Dec. 2024, doi: 10.1186/S40537-024-01002-8.
- [69] P. Jafary, D. Shojaei, A. Rajabifard, and T. Ngo, “Automated land valuation models: A comparative study of four machine learning and deep learning methods based on a comprehensive range of influential factors,” *Cities*, vol. 151, p. 105115, Aug. 2024, doi: 10.1016/J.CITIES.2024.105115.
- [70] A. Gavros, Y. C. Hsu, and K. Karatzas, “Modelling Smell Events in Urban Pittsburgh with Machine and Deep Learning Techniques,” *Atmosphere (Basel)*, vol. 15, no. 6, Jun. 2024, doi: 10.3390/ATMOS15060731.
- [71] A. Masood *et al.*, “Improving PM2.5 prediction in New Delhi using a hybrid extreme learning machine coupled with snake optimization algorithm,” *Sci. Rep.*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/S41598-023-47492-Z.
- [72] A. A. Khadom, S. Albawi, A. J. Abboud, H. B. Mahood, and Q. Hassan, “Predicting air quality index and fine particulate matter levels in Bagdad city using advanced machine learning and deep learning techniques,” *J. Atmos. Solar-Terrestrial Phys.*, vol. 262, p. 106312, 2024, doi: <https://doi.org/10.1016/j.jastp.2024.106312>.
- [73] Q. Liu, B. Cui, and Z. Liu, “Air Quality Class Prediction Using Machine Learning Methods Based on Monitoring Data and Secondary Modeling,” *Atmosphere (Basel)*, vol. 15, no. 5, 2024, doi: 10.3390/atmos15050553.
- [74] T. M. T. Lei, J. Cai, A. H. Molla, T. A. Kurniawan, and S. S.-K. Kong, “Evaluation of Machine Learning Models in Air Pollution Prediction for a Case Study of Macau as an Effort to Comply with UN Sustainable Development Goals,” *Sustainability*, vol. 16, no. 17, p. 7477, Aug. 2024, doi: 10.3390/SU16177477.
- [75] C. Shetty *et al.*, “A Machine Learning Approach for Environmental Assessment on Air Quality and Mitigation Strategy,” *J. Eng.*, vol. 2024, pp. 1–16, 2024, [Online]. Available: <https://research.ebsco.com/c/tjf7vj/search/details/trzvwsqlz?db=egs%2Cfap%2Caph&limiters=None&q=A Machine Learning Approach for Environmental Assessment on Air Quality and Mitigation Strategy>.
- [76] G. S. Bhat *et al.*, “Machine Learning-Based Asthma Risk Prediction Using IoT and Smartphone Applications,” *IEEE Access*, vol. 9, pp. 118708–118715, 2021, doi: 10.1109/ACCESS.2021.3103897.
- [77] H. Liao, L. Yuan, M. Wu, and H. Chen, “Air quality prediction by integrating mechanism model and machine learning model,” *Sci. Total Environ.*, vol. 899, Nov. 2023, doi: 10.1016/J.SCITOTENV.2023.165646.
- [78] M. A. Khan, H.-C. Kim, and H. Park, “Leveraging Machine Learning for Fault-Tolerant Air Pollutants Monitoring for a Smart City Design,” *Electron.*, vol. 11, no. 19, 2022, doi: 10.3390/electronics11193122.
- [79] A. K. Dutta *et al.*, “Hybrid Deep Learning Enabled Air Pollution Monitoring in ITS Environment,” *Comput. Mater. Contin.*, vol. 72, no. 1, pp. 1157 – 1172, 2022, doi: 10.32604/cmc.2022.024109.
- [80] K. Alnowaiser, A. A. Alarfaj, E. A. Alabdulqader, M. Umer, L. Cascone, and B. Alankar, “IoT based smart framework to predict air quality in congested traffic areas using SV-CNN ensemble and KNN imputation model,” *Comput. Electr. Eng.*, vol. 118, 2024, doi: 10.1016/j.compeleceng.2024.109311.