

# EDUCACIÓN

SECRETARÍA DE EDUCACIÓN PÚBLICA



Instituto Politécnico Nacional  
"La Técnica al Servicio de la Patria"

# Research in Computing Science

**Vol. 155 No. 4**  
**April 2026**



# Research in Computing Science

---

## Series Editorial Board

### Editors-in-Chief:

*Grigori Sidorov, CIC-IPN, Mexico*  
*Gerhard X. Ritter, University of Florida, USA*  
*Jean Serra, Ecole des Mines de Paris, France*  
*Ulises Cortés, UPC, Barcelona, Spain*

### Associate Editors:

*Jesús Angulo, Ecole des Mines de Paris, France*  
*Jihad El-Sana, Ben-Gurion Univ. of the Negev, Israel*  
*Alexander Gelbukh, CIC-IPN, Mexico*  
*Ioannis Kakadiaris, University of Houston, USA*  
*Petros Maragos, Nat. Tech. Univ. of Athens, Greece*  
*Julian Padget, University of Bath, UK*  
*Mateo Valero, UPC, Barcelona, Spain*  
*Olga Kolesnikova, ESCOM-IPN, Mexico*  
*Rafael Guzmán, Univ. of Guanajuato, Mexico*  
*Juan Manuel Torres Moreno, U. of Avignon, France*  
*Miguel González-Mendoza, ITESM, Mexico*

### Editorial Coordination:

*Alejandra Ramos Porras*

**Research in Computing Science**, Año 25, Volumen 155, No. 4, abril de 2026, es una publicación mensual, editada por el Instituto Politécnico Nacional, a través del Centro de Investigación en Computación. Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othon de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, Ciudad de México, Tel. 57 29 60 00, ext. 56571. <https://www.rcs.cic.ipn.mx>. Editor responsable: Dr. Grigori Sidorov. Reserva de Derechos al Uso Exclusivo del Título No. 04-2019-082310242100-203. ISSN: en trámite, ambos otorgados por el Instituto Politécnico Nacional de Derecho de Autor. Responsable de la última actualización de este número: el Centro de Investigación en Computación, Dr. Grigori Sidorov, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othon de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738. Fecha de última modificación 01 de abril de 2026.

Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor de la publicación.

Queda estrictamente prohibida la reproducción total o parcial de los contenidos e imágenes de la publicación sin previa autorización del Instituto Politécnico Nacional.

**Research in Computing Science**, year 25, Volume 155, No. 4, April 2026, is published monthly by the Center for Computing Research of IPN.

The opinions expressed by the authors does not necessarily reflect the editor's posture.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of Centre for Computing Research of the IPN.

# Advances in Artificial Intelligence

Leonardo Morales (ed.)



Instituto Politécnico Nacional  
"La Técnica al Servicio de la Patria"



Instituto Politécnico Nacional, Centro de Investigación en Computación  
México 2026

## ISSN: in process

---

Copyright © Instituto Politécnico Nacional 2026  
Formerly ISSNs: 1870-4069, 1665-9899

Instituto Politécnico Nacional (IPN)  
Centro de Investigación en Computación (CIC)  
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal  
Unidad Profesional “Adolfo López Mateos”, Zacatenco  
07738, México D.F., México

<http://www.rcs.cic.ipn.mx>

<http://www.ipn.mx>

<http://www.cic.ipn.mx>

The editors and the publisher of this journal have made their best effort in preparing this special issue, but make no warranty of any kind, expressed or implied, with regard to the information contained in this volume.

All rights reserved. No part of this publication may be reproduced, stored on a retrieval system or transmitted, in any form or by any means, including electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the Instituto Politécnico Nacional, except for personal or classroom use provided that copies bear the full citation notice provided on the first page of each paper.

Indexed in LATINDEX, DBLP and Periodica

Electronic edition

## Table of Contents

	Page
Uso de SVM para detección de fallas en una unidad controladora de luz LED .....	5
<i>Jorge L. Gutiérrez Palomino, Óscar Hernández-Uribe</i>	
Perfeccionamientos en estrategias de marketing utilizando inteligencia artificial para un creador de contenido.....	11
<i>Blanca Cecilia López-Ramírez, Mariana Rojas Delgado, José Luis Camargo Orduño, Baltazar García Alcocer</i>	
Predicción de la tasa de degradación de un proceso de fotocatalisis heterogénea: comparativa entre una red neuronal artificial y un modelo de bosques aleatorios.....	17
<i>Katia L. Rocha-Aguilar, José I. Hernández-Vega, Cynthia G. Aba-Guevara, Alejandro Alonzo-García, René Sanjuan-Galindo</i>	



## Uso de SVM para detección de fallas en una unidad controladora de luz LED

Jorge L. Gutiérrez Palomino, Óscar Hernández-Uribe

Centro de Tecnología Avanzada,  
Posgrado CIATEQ A.C.,  
México

jorgeluis.gutierrezpalomino@gmail.com,  
oscar.hernandez@ciateq.mx

**Resumen.** Los algoritmos de aprendizaje automático supervisado son cada vez más usados utilizados para resolver problemas de clasificación y predicción en la industria automotriz. En el desarrollo de sistemas de iluminación para el automóvil, como lo son las Unidades Controladoras de luz LED es común desarrollar algoritmos para la detección de fallas de sus propios sensores y/o dispositivos internos; por lo que su detección depende de una correcta caracterización del algoritmo. Por lo que el presente experimento hace uso de los algoritmos de clasificación máquinas de vectores de soporte y regresión logística para el modelado del algoritmo de detección de fallas del tipo circuito abierto en una matriz LED. Dentro de los resultados obtenidos, se identifican las variables con mayor relación para la detección de la falla anteriormente mencionada, así como su exactitud y precisión en escenarios de atenuación de luminosidad donde el algoritmo presenta vulnerabilidad.

**Palabras clave:** Aprendizaje automático, falla circuito abierto, controlador LED.

### Using SVM for Fault Detection in a LED Light Controller Unit

**Abstract.** Supervised machine learning algorithms are increasingly used to solve classification and prediction problems in the automotive industry. In the development of automotive lighting systems, such as LED Light Controller Units, it is common to develop algorithms for detecting faults in their own sensors and/or internal devices; therefore, their detection depends on a correct characterization of the algorithm. This experiment uses support vector machine and logistic regression classification algorithms to model an open-circuit fault detection algorithm in an LED array. The results identify the variables most closely related to the detection of the aforementioned fault, as well as its accuracy and precision in dimming scenarios where the algorithm exhibits vulnerability.

**Keywords:** Machine learning, open-circuit fault, LED controller.

## 1. Introducción

La industria automotriz está creando sistemas eléctricos y electrónicos cada vez más complejos, que deben cumplir con normas para garantizar su funcionamiento y la seguridad de las personas [1]. Además, se está enfocando en la eficiencia y reducción de costos integrando soluciones con el internet de las cosas (IoT) y la inteligencia artificial [2]. Por ejemplo, Kumar y Jain [3] usan máquinas de vectores de soporte (SVM), para clasificar el comportamiento de conducción de un conductor (consumo de combustible, patrones de frenado, estabilidad de dirección y velocidad). Gong et al. [4] proponen un modelo de aprendizaje profundo (DL) aplicado a la predicción de fallas, tales como funcionamiento anormal del motor. Abdulkareem et al. [5] presentan un análisis usando SVM para la detección de fallos en los sensores de velocidad en vehículos. Safavi et al. [6] utilizan dos arquitecturas de DL y SVM, y Biddle y Fallah [7] usan solo SVM, ambos se enfocan en la detección, aislamiento e identificación de fallas en sistemas multisensores, tales como fallas erráticas en la presión de frenado y ángulo del volante. Kaplan et al. [8] muestra un modelo DL de diagnóstico de fallas en motores de inducción, tales como cortocircuito y circuito abierto.

En México, Continental Automotive cuenta con 21 locaciones de las cuales 3 son centros de desarrollo especializados en el diseño de productos electrónicos para el automóvil [9]. Dentro de su portafolio de productos cuenta con unidades controladoras de luz LED, la cuales son capaces de controlar tiras de LED con voltajes de hasta 75V [10]. La tecnología LED ha tenido gran aceptación en la industria automotriz, brindando mayor confiabilidad y un menor consumo de energía [11]. Sin embargo, los LED son dispositivos semiconductores de unión tipo p-n, donde sus propiedades eléctricas y ópticas dependen de la temperatura, por lo que solo el 30% de la energía sirve para emitir luz [12], ocasionando una degradación eventual en su vida útil y en el rendimiento de iluminación. Para evitarlo y poder generar un funcionamiento estable son necesarias técnicas para el manejo de temperatura [13]. Por lo anterior, se han impulsado soluciones desde distintas perspectivas, tales como el uso de disipadores de calor metálicos y materiales térmicos en el PCB [14,15]. Desde la perspectiva de software como área de investigación, algunos métodos comunes de atenuación son utilizados tales como TRIAC, PWM, PAM y control de corriente [16]. En este artículo, PWM y control de corriente son utilizados haciendo uso de SVM y regresión logística para la detección de fallas en una unidad controladora de luz LED.

## 2. Metodología

La propuesta incluye el uso de algoritmos de clasificación SVM y regresión logística, para caracterizar de manera efectiva la detección de fallas del tipo circuito abierto en una matriz LED conectada a un sistema embebido llamado unidad controladora de luz LED (UCLED). El proceso de entrenamiento y validación se realizó utilizando una técnica de validación cruzada, que consistió dividir el conjunto de datos (DS) en dos partes. Para ambos algoritmos un 80% de las muestras se usaron para entrenar el modelo y el restante se reservó para validarlo. La Figura 1, describe el

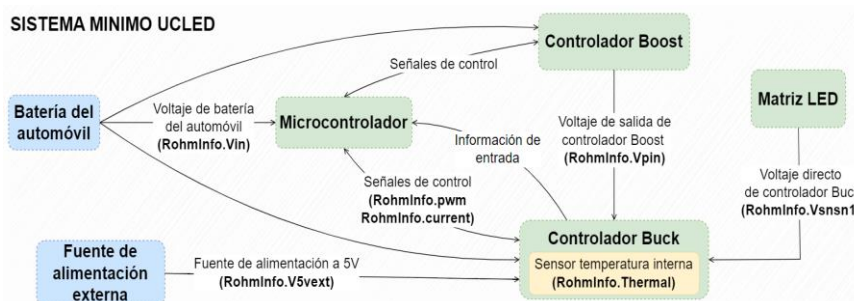


Fig. 1. Sistema mínimo de la unidad controladora LED.

Tabla 1. Conjuntos de datos y variables.

ID	Cantidad LEDs	Corriente mA	PWM %
1	1	150	5
2	7	750	100
3	12	1500	100

sistema mínimo de UCLED, esta se compone de los siguientes elementos: Batería del automóvil; Fuente de alimentación externa de 5 voltios para el controlador Buck; microcontrolador para enviar y recibir información de los controladores; controlador Boost para regular el voltaje de entrada del controlador Buck; controlador Buck, controlador regulador de corriente y PWM de la matriz LED; y la matriz LED, que es la salida.

El controlador Buck a utilizar es el BD18398RUV-M con su circuito de aplicación típico [17]. El microcontrolador de la UCLED envía parámetros de corriente y PWM, recibirá información de voltajes y temperatura por parte del controlador Buck. Con esta información se caracterizó el detector de fallas del tipo circuito abierto en la matriz LED, usando SVM y regresión logística. Programando esto en Python, empleando las bibliotecas de Pandas, Numpy y Scikit-learn, y manejando el DS en formato CSV. Los DS utilizados fueron obtenidos mediante el uso de la herramienta daqIDEA [18]. La respuesta esperada a la salida es verdadero o falso si hay o no una falla del tipo circuito abierto en la matriz LED. En las pruebas se utilizaron 17 DS, induciendo una falla del tipo circuito abierto en la matriz LED, reportamos los resultados de 3 DS. Los DS fueron creados acorde a las siguientes características: modulación PWM, número de leds utilizados y corriente en el controlador Buck.

La Tabla 1 muestra datos cuantitativos de los DS. Cada DS cuenta con alrededor de 500 a 1000 muestras. Las variables de entrada almacenadas en los DS son las siguientes: RohmInfo.current, valor de corriente en el controlador Buck (mA); RohmInfo.pwm, valor de PWM en el controlador Buck (%); RohmInfo.Thermal, valor de temperatura interna del controlador Buck (°C); RohmInfo.V5vext, valor de voltaje externo medido por el controlador Buck (mV); RohmInfo.Vin, valor de voltaje de

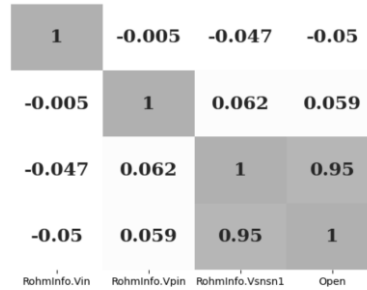


Fig. 2. Correlación de variables de DS.

Tabla 2. Resultados exactitud y precisión.

Algoritmo	Configuración	Exactitud	Precisión
SVM	1 led, 150 mA, 5% PWM	0.9889	0.9829
SVM	7 leds, 150 mA, 5% PWM	0.9893	0.9736
SVM	12 leds, 150 mA, 5% PWM	0.9811	0.9898
RegLog	1 led, 150 mA, 5% PWM	0.9981	0.9971
RegLog	7 leds, 150 mA, 5% PWM	0.9936	0.9844
RegLog	12 leds, 150 mA, 5% PWM	0.9641	0.9144

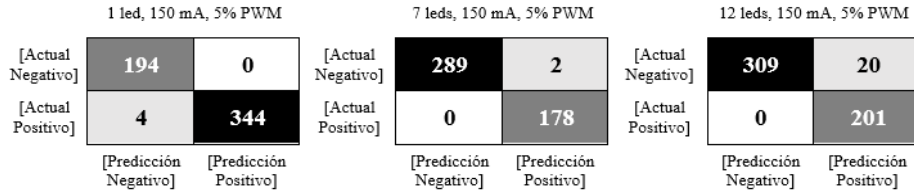


Fig. 3. Matrices de confusión.

batería medido por el controlador Buck (mV); RohmInfo.Vpin, valor de voltaje de salida del controlador Boost (mV); RohmInfo.Vsnsn1, valor de voltaje consumido por la matriz LED (mV).

Las variables de salida almacenadas en los DS son las siguientes: Voltage\_diff\_out, valor de diferencia de voltaje entre RohmInfo.Vpin y RohmInfo.Vsnsn1; OL\_Detect, variable que contiene valor verdadero o falso para la detección de falla del tipo circuito abierto en la matriz LED. Finalmente, se obtienen las métricas de exactitud, precisión y matriz de confusión para la evaluación del modelo.

### 3. Resultados y conclusiones

Mediante el uso de correlaciones, se observa que las variables relacionadas a la variable de salida son: RohmInfo.Vsnsn1 y RohmInfo.Vpin como se muestra en la

figura 2. Dado que la variable RohmInfo.Vpin tiene un coeficiente no tan cercano a 1, se decide no utilizarla para el entrenamiento de los modelos. Por lo tanto, la variable “x” para el entrenamiento es RohmInfo.Vsnsn1, y la variable “y” es OL\_Detect. Los resultados obtenidos de exactitud y precisión son mostrados en la tabla 2; mientras que las matrices de confusión son mostradas en la figura 3.

Como conclusión, se deduce que la variable RohmInfo.Vsnsn1 es más que suficiente para crear una estrategia en base a voltajes para la detección de fallas de circuito abierto en la matriz LED. Al haber una falla, el valor de RohmInfo.Vsnsn1 tiende a acercarse al valor de RohmInfo.Vpin, por lo que un valor de umbral finamente ajustado entre ambos valores permite una correcta detección de la falla. Esta solución por software podría minimizar gastos en el uso de dispositivos más complejos para realizar la misma función.

## Referencias

1. Damsgaard, H.J., Grenier, A., Katare, D., Taufique, Z., Shakibhamedan, S., Troccoli, T., Chatzitsompanis, G., Kanduri, A., Ometov, A., Ding, A.Y., Taherinejad, N., Karakonstantis, G., Woods, R., Nurmi, J.: Adaptive approximate computing in edge AI and IoT applications: A review. *Journal of Systems Architecture* 150, 103114 (2024). DOI: 10.1016/j.sysarc.2024.103114.
2. Ullah, I., Adhikari, D., Su, X., Palmieri, F., Wu, C., Choi, C.: Integration of data science with the intelligent IoT (IIoT): Current challenges and future perspectives. *Digital Communications and Networks* 11(2), 280–298 (2025). DOI: 10.1016/j.dcan.2024.02.007.
3. Kumar, R., Jain, A.: Driving behavior analysis and classification by vehicle OBD data using machine learning. *The Journal of Supercomputing* 79(16), 18800–18819 (2023). DOI: 10.1007/s11227-023-05364-3.
4. Gong, C.S.A., Su, C.H.S., Chen, Y.H., Guu, D.Y.: How to implement automotive fault diagnosis using artificial intelligence scheme. *Micromachines* 13(9), 1380 (2022). DOI: 10.3390/mi13091380.
5. Abdulkareem, A.Q., Humod, A.T., Ahmed, O.A.: Robust pattern recognition based fault detection and isolation method for ABS speed sensor. *International Journal of Automotive Technology* 23(6), 1747–1754 (2022). DOI: 10.1007/s12239-022-0152-5.
6. Safavi, S., Safavi, M.A., Hamid, H., Fallah, S.: Multi-sensor fault detection, identification, isolation and health forecasting for autonomous vehicles. *Sensors* 21(7), 2547 (2021). DOI: 10.3390/s21072547.
7. Biddle, L., Fallah, S.: A novel fault detection, identification and prediction approach for autonomous vehicle controllers using SVM. *Automotive Innovation* 4(3), 301–314 (2021). DOI: 10.1007/s42154-021-00138-0.
8. Kaplan, H., Tehrani, K., Jamshidi, M.: A fault diagnosis design based on deep learning approach for electric vehicle applications. *Energies* 14(20), 6599 (2021). DOI: 10.3390/en14206599.
9. Continental AG: Continental Mexico. <https://www.continental.com/es-mx/compania/continental-mexico/>
10. Continental AG: Continental focuses on LED innovations. <https://www.continental.com/en/press/press-releases/2014-02-28-full-led/>
11. Yong, P.C.: Led lamps replace incandescent bulbs in automotive applications. *ATZautotechnology* 1(6), 56–57 (2001). DOI: 10.1007/BF03246657.

12. Jeong, S.-H., Kim, J.-Y., Xiao, X., Kim, Y.-S.: Junction temperature control for the automotive LED lamp based on the current feed forward strategy. *Optik* 232, 166535 (2021). DOI: 10.1016/j.ijleo.2021.166535.
13. Wang, J., Cai, Y.X., Li, X.H., Zhao, X.J., Zhang, C.: Design of automotive headlamp with high-power LEDs. *International Journal of Automotive Technology* 15(4), 673–681 (2014). DOI: 10.1007/s12239-014-0070-2.
14. Huang, D.-S., Chen, T.-C., Tsai, L.-T., Lin, M.-T.: Design of fins with a grooved heat pipe for dissipation of heat from high-powered automotive LED headlights. *Energy Conversion and Management* 180, 550–558 (2019). DOI: 10.1016/j.enconman.2018.11.021.
15. Sevilgen, G., Kiliç, M., Aktaş, M.: Dual-separated cooling channel performance evaluation for high-power LED PCB in automotive headlight. *Case Studies in Thermal Engineering* 25, 100985 (2021). DOI: 10.1016/j.csite.2021.100985.
16. Hsia, S.-C., Chen, K.-C.: Universal dimmable driving system for various power LEDs using digital adaptation. *Electrical Engineering* 105, 151–161 (2023). DOI: 10.1007/s00202-022-01655-9.
17. ROHM Co., Ltd.: BD18398RUV-M 3ch current LED driver with SPI for automotive. <https://www.rohm.com/products/power-management/led-drivers/step-down-buck/bd18398ruv-m-product>
18. iSYSTEM: daqIDEA - Visualization Tool. <https://www.isystem.com/products/software/daqidea.html>

## **Perfeccionamientos en estrategias de marketing utilizando inteligencia artificial para un creador de contenido**

Blanca Cecilia López-Ramírez, Mariana Rojas Delgado,  
José Luis Camargo Orduño, Baltazar García Alcocer

Tecnológico Nacional de México/I. T. Roque, Celaya, Guanajuato,  
México

{blanca.lr, jose.co, mariana.rd, 120980600}@roque.tecnm.mx

**Resumen.** Actualmente el posicionamiento de marca, juega un papel fundamental dentro de las estrategias de negocios ya que esta representa el lugar que se ocupa en la mente del consumidor, obteniendo así la diferencia entre sus competidores. El uso de estrategias de posicionamiento, representa un papel clave para las organizaciones, las cuales han optado por el uso en conjunto de técnicas tradicionales de marketing, así como el de herramientas de Inteligencia Artificial (IA) que tienen un impacto significativo, ofrecen ventajas y mejoras en diversas facetas del marketing digital, estas aplicaciones, no solo permiten mejorar la eficiencia y la efectividad de las estrategias de marketing, también ofrecen nuevas posibilidades para entender e interactuar con los clientes. Se presenta el caso de uso enfocado al benchmarking para un creador de contenido para su marca personal, a través de un diagnóstico realizado con preguntas estratégicas elaboradas con el acompañamiento de herramientas de IA y con la adopción de la estrategia propuesta, se lograron identificar las necesidades latentes para la marca personal, así como la detección de necesidades de cambio de contenido y calidad, lo cual implicó una mejor respuesta del mercado meta teniendo como resultados un nuevo número de suscriptores, así como una mayor permanencia.

**Palabras clave:** Gestión de marca, tecnologías digitales, grafos, branding.

### **Refinements in Marketing Strategies Using Artificial Intelligence for a Content Creator**

**Abstract.** Brand positioning currently plays a fundamental role in business strategies, as it represents the place a brand occupies in the consumer's mind, thus differentiating it from its competitors. The use of positioning strategies is key for organizations, which have opted for a combination of traditional marketing techniques and Artificial Intelligence (AI) tools. These tools have a significant impact, offering advantages and improvements in various facets of digital marketing. These applications not only improve the efficiency and effectiveness of marketing strategies but also offer new possibilities for understanding and

interacting with customers. This paper presents a benchmarking case study for a content creator's personal brand. Through a diagnostic process using strategic questions developed with the support of AI tools, and with the adoption of the proposed strategy, latent needs for the personal brand were identified, as well as the need for changes in content and quality. This resulted in a better response from the target market, leading to a new number of subscribers and increased engagement.

**Keywords:** Brand management, digital technologies, graphs, branding.

## 1. Introducción

Desde los inicios del marketing, las empresas comenzaron a reconocer la importancia de entender y las satisfacer las necesidades y deseos de los consumidores para pro-mover sus productos. Hoy en día es importante superponer la Inteligencia Artificial en los diferentes procesos del marketing. Un problema para las Micro, Pequeñas y Medianas Empresas (MiPyMES), marcas personales e instituciones, es que la cultura del uso de estrategias de marketing para posicionar sus marcas, no existe como prioridad [1]. El marketing bien elaborado concede la colocación de un producto o servicio y asegura la escalabilidad de su Mercado en un futuro [2]. Por ello, la importancia de estrategias adaptadas que aseguren resultados favorables.

La participación del IA en el marketing ha experimentado un crecimiento exponencial permitiendo construir una marca, incrementar las ventas y ganar nuevos mercados.

La digitalización fue el empuje para las MiPyMES en México antes y durante la pandemia. Ramos en el 2021 [3], en su nota menciona que el 81% de los propietarios prefieren vender en línea y que la digitalización es un camino crucial en el entorno cambiante de hoy en día. Una empresa que no cuenta con procesos robustos, y no invierte en marketing disminuyen la esperanza de vida [4]. En contraste, grandes empresas como Starbucks®, ha utilizado diferentes estrategias de marketing con el fin de crear una marca sólida para incrementar su volumen de ventas, y con el uso de branding lograr la consistencia de la marca, la comunicación social, programas de lealtad, tecnología, marketing experiencial e innovación del producto [5].

De lo anterior que, aunque existen etapas o procesos estratégicos de marketing tradicional [6], la razón de implementar una metodología para marketing digital con el uso de herramientas con inteligencia artificial en cada una de las diferentes etapas, propiciará el impacto de los resultados significativos, que va desde obtener la marca del cliente hasta la evaluación de los resultados de la estrategia implementada.

## 2. Materiales y métodos

Las fases que fueron propuestas en el caso de estudio de una marca personal conocido como creador de contenido se muestran en el esquema de la Figura 1. Por lo tanto, la fase de diagnóstico se llevó a cabo con el desarrollo de la entrevista, que fue materializada con el modelo predictivo de lenguaje generativo (chat-gpt v 3.5); Considerando las etapas del diagnóstico de la empresa, el análisis del entorno, la

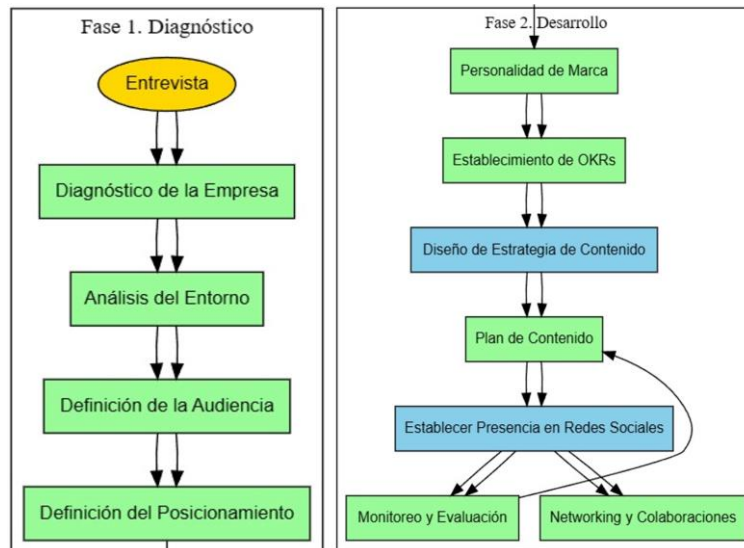


Fig. 1. Fases y etapas que fueron propuestas en el caso de estudio de un creador de contenido.

definición de la audiencia, así como la definición del posicionamiento. La herramienta inteligente sugirió una encuesta clasificada en 6 secciones, con el propósito de estudiar la estrategia personal, la comunicación, las colaboraciones y networking, el desarrollo profesional y la interacción con la audiencia.

Por otro lado, en la fase 2 se observó que la personalidad de la marca fue propuesta en la construcción del branding con la intención de crear una imagen que pueda impactar en los clientes.

Posteriormente, se fijaron Objetivos y Resultados Clave (OKR's por sus siglas en inglés), que son: a) cantidad de suscriptores por mes, b) tiempo promedio de permanencia, c) número de suscriptores nuevos y, d) cantidad de vistas del video completo, éstos aplicados a cada uno de los medios de comunicación, YouTube, Tiktok, Instagram y Facebook. En esta etapa se obtuvo el logo, la marca, la estrategia de mercado, el estado del arte y, la definición de los medios de difusión.

Los números de colaboraciones confirmadas, el alcance ampliado, la participación en grupos/comunidades y el feedback de colaboradores son indicadores clave de la etapa de colaboración y networking. El monitoreo de los OKR's se diseñó con la finalidad de identificar oportunidades de mejora en el momento y plantear acciones inmediatas, de tal manera que se registró diariamente, sin embargo, el resultado final del indicador se cierra cada mes.

### 3. Resultados

Entre las preguntas de mayor impacto dentro de la entrevista son:

- a) ¿Cómo aprovechas las colaboraciones y el networking para fortalecer tu marca personal?, b) ¿Qué plataformas digitales prefieres para promover tu marca personal y

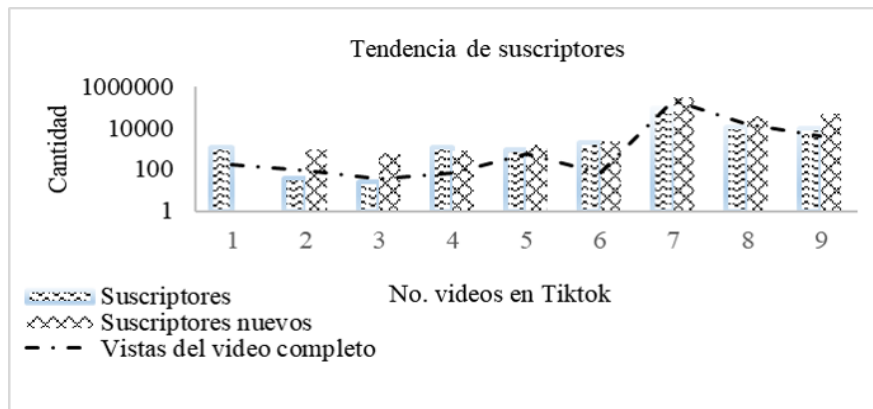


Fig. 2. Comportamiento de 3 OKR's.

por qué?, c) ¿Cómo fomentas la interacción y participación de la audiencia en los contenidos?, d) ¿Has recibido feedback directo de tu audiencia y cómo lo gestionas?

El bechmarking del creador de contenido para su marca personal enfocado a videojuegos, más específicamente en shooters, fue propuesta por el cliente y se denomina T3nNgu y, cuyo propósito es, que la identificación del logo llegue a provocar el branding en el espectador o suscriptor. En cuanto al establecimiento de contenido, las metas con enfoque principal para Tiktok se presentan en la Figura 2.

Derivado de la evaluación del mercado se obtuvieron resultados deseables, se adquirió el aprendizaje de los gustos de los seguidores, uso de las herramientas de IA, y se identificó la necesidad de cambiar el contenido y la calidad. Además, se visualizó una mejor respuesta en comparación con los videos anteriormente publicados, aunque el tiempo del video en promedio, era el mismo (9 s). Los videos publicados por trimestre (escala logarítmica base 10, en el eje Y) se muestran en la Figura 2, donde se cuenta con suscriptores nuevos en todas las visitas, y la línea punteada marca el comportamiento de la vista completa de los videos.

#### 4. Conclusiones y trabajo a futuro

Para el caso estudiado, en promedio la retención de los suscriptores aumentó y se habrá de reducir el contenido a menor cantidad de tiempo. Con respecto a la metodología, la herramienta de IA sugiere recomendaciones para alcanzar un mayor número de seguidores como son: encuestas y desafíos para retener la atención del suscriptor. Además, dependiendo de la marca o de la empresa la metodología es adaptable a las necesidades.

#### Referencias

- Gilligan, C., Wilson, R.M.S.: Strategic marketing planning. Butterworth-Heinemann, Oxford, Reino Unido (2003). DOI: 10.4324/9780080481142.

2. Kerzner, H.: Project management: a systems approach to planning, scheduling, and controlling. 12th edn. John Wiley & Sons, Hoboken, NJ (2017)
3. Liberos Hoppe, E., Ahumada Luyando, S., Sánchez Ahumada, M.: Inteligencia artificial para el marketing: cómo la tecnología revolucionará tu estrategia. 1st edn. ESIC Editorial, Madrid, España (2024)
4. Ramos, M.: 45% de las pymes digitales en México se crearon en los últimos dos años [Goddady]. Marketing4eCommerce México (2021), <https://marketing4ecommerce.mx/45-de-las-pymes-digitales-en-mexico-se-crearon-en-los-ultimos-dos-anos-goddady/>
5. Ruiz, J.P.: Porcentaje de empresas que fracasan en México. Factura-e (2024), <https://www.factura-e.mx/blog/porcentaje-de-empresas-que-fracasan-en-mexico/>
6. StartupTalky: The marketing strategies that made Starbucks a global phenomenon. StartupTalky, <https://startuptalky.com/starbucks-marketing-strategies/>



# **Predicción de la tasa de degradación de un proceso de fotocatalisis heterogénea: Comparativa entre una red neuronal artificial y un modelo de bosques aleatorios**

Katia L. Rocha-Aguilar, José I. Hernández-Vega,  
Cinthia G. Aba-Guevara, Alejandro Alonzo-García,  
René Sanjuan-Galindo

Instituto Tecnológico de Nuevo León,  
División de Estudios de Posgrado e Investigación,  
México

{mg15480487, jose.hv, cinthia.ag, alejandro.ag,  
rene.sg}@nuevoleon.tecnm.mx

**Resumen.** Dada la creciente demanda del consumo de agua, la escasez de este recurso, y el aumento de las aguas residuales generadas por la actividad humana, se han realizado esfuerzos para optimizar, mejorar y proponer tratamientos de aguas residuales alternativos, entre los cuales se encuentra la fotocatalisis heterogénea, un proceso de oxidación avanzado. Varios estudios se han llevado a cabo en un esfuerzo por lograr su implementación a gran escala, sin embargo, el análisis de la eficiencia del proceso es complejo, ya que en el proceso de tratamiento de aguas residuales intervienen diversos factores que resultan complejos de modelar, predecir y/o controlar, y se requiere la realización de un gran número de experimentos. Este artículo propone la implementación de modelos de inteligencia artificial para la predicción de la tasa de degradación fotocatalítica (información relevante para el estudio de la eficiencia del proceso) haciendo uso de TiO<sub>2</sub> como catalizador, realizando una comparativa entre un modelo de redes neuronales artificiales y uno de bosques aleatorios.

**Palabras clave:** Inteligencia artificial, fotocatalisis heterogénea, aprendizaje automático.

## **Prediction of the Degradation Rate of a Heterogeneous Photocatalysis Process: Comparison between an Artificial Neural Network and a Random Forest Model**

**Abstract.** Given the growing demand for water consumption, the scarcity of this resource, and the increase in wastewater generated by human activity, efforts have been made to optimize, improve, and propose alternative wastewater treatments, including heterogeneous photocatalysis, an advanced oxidation

process. Several studies have been conducted in an effort to achieve its large-scale implementation; however, analyzing the efficiency of the process is complex, since the wastewater treatment process involves various factors that are difficult to model, predict, and/or control, and requires a large number of experiments. This article proposes the implementation of artificial intelligence models to predict the photocatalytic degradation rate (relevant information for studying the efficiency of the process) using TiO<sub>2</sub> as a catalyst, comparing an artificial neural network model with a random forest model.

**Keywords:** Artificial intelligence, heterogeneous photocatalysis, machine learning.

## 1. Introducción

El aumento del consumo del recurso hídrico para las diferentes actividades humanas representa una mayor generación de aguas residuales, por lo que el impacto ambiental no solo se ve reflejado en la sobre-extracción de los mantos acuíferos, sino también en la contaminación de los ecosistemas debido a los contaminantes vertidos en las corrientes naturales [1]. Es por lo anterior que existen procesos de tratamiento de estas aguas contaminadas, con el fin de reutilizarlas y reducir el impacto ambiental. Representan un riesgo para el medioambiente y la salud la presencia de contaminantes orgánicos en aguas residuales, muchos de los cuales son difíciles de remover o eliminar mediante métodos convencionales de tratamiento o procesos de degradación natural. La degradación por fotocatalisis es una alternativa que ha mostrado poseer un gran potencial en la degradación de estos compuestos, llegando incluso a lograr su completa mineralización [2]. El dióxido de titanio (TiO<sub>2</sub>) como fotocatalizador, es uno de los óxidos metálicos ampliamente estudiados por su actividad fotocatalítica y sus propiedades, además de tratarse de un compuesto estable y de barata adquisición [3], sin embargo, su implementación en procesos de fotocatalisis heterogénea presenta varios retos para su implementación a gran escala, entre los cuales se encuentran el estudio de la cinética de degradación y la aplicabilidad para la degradación de diversos contaminantes [4].

El conocimiento sobre las velocidades de reacción bajo determinadas condiciones es esencial para la optimización del proceso, las cuales son analizadas mediante numerosos ensayos experimentales, sin embargo, parámetros que pueden ser óptimos para un determinado tipo de contaminante no son los mismos para otro. Son diversos los factores que impactan en el desempeño de este proceso, como lo son la concentración y tipo de contaminante, la concentración del catalizador, el pH, la fuente de luz, entre otros. Los modelos de inteligencia artificial (IA) se han implementado en diferentes ámbitos de los procesos de tratamientos de aguas residuales, para abordar tareas de optimización, predicción, control y modelado de aguas residuales [5]. Como una subcategoría de estos modelos, se han desarrollado algoritmos de aprendizaje automático con la capacidad de encontrar patrones presentes en los datos y adaptarse al comportamiento no lineal, por ejemplo, de la degradación de contaminantes, realizando esta predicción con base en los parámetros que intervienen en el proceso. En la

**Tabla 1.** Consideraciones para la selección de los datos.

<b>Tipo de reactor</b>	Reactor Batch	Debido a que las variaciones físicas del diseño en comparación con otros son menores debido a su sencillez, reduciendo sesgos que estas puedan tener en el modelo.
<b>Tipo de fotocatalizador</b>	TiO <sub>2</sub>	Debido a su accesibilidad, bajo costo y por ser uno de los más utilizados y estudiados debido a sus propiedades fisicoquímicas.
<b>Otros parámetros</b>	<u>Tipo de luz:</u> UV	Tipo de fotocatalizador: Nanopartículas TiO <sub>2</sub> – P25 en suspensión

**Tabla 2.** Estructura de la BD.

Columnas	Rango	Unidad
Nombre del contaminante	-	-
Intensidad de la luz	0.176 - 75	mW/cm <sup>2</sup>
Temperatura	20 – 60	°C
Concentración del contaminante	0.13 – 342.47	mg/L
Concentración del catalizador	0.001 – 7.5	g/L
pH inicial	2 – 11	-
pKa ácido	-3 – 15.96	-
pKa básico	-1.99 – 9.27	-
Longitud de onda	172 - 400	nm
Potencia de la lámpara	6 - 1500	W

degradación mediante fotocatalisis heterogénea, resulta complejo modelar y predecir el comportamiento y eficiencia del proceso, dónde además es requerida para su análisis la realización de un gran número de experimentos, donde se evalúa el impacto de las variables que intervienen en él.

Por lo anterior, se presentan en este artículo los resultados de 2 propuesta de modelo de aprendizaje automático, utilizando redes neuronales artificiales (RNA) y bosques aleatorios (BA), para la predicción de la tasa de degradación de contaminantes bajo determinados parámetros experimentales.

## 2. Metodología

Una vez identificados y delimitados los parámetros que impactan en la eficiencia del proceso, se procedió con la adquisición de los datos que conforman la base de datos, la cual fue utilizada para el entrenamiento y validación de los modelos de aprendizaje automático. La información fue extraída de los resultados experimentales reportados en la literatura, artículos científicos especializados.

En la selección de los artículos se consideraron aquellos que cumplieran con los aspectos en la metodología de la experimentación descritos en la Tabla 1, con el fin de

**Tabla 3.** Análisis exploratorio de los datos.

Parámetro	Correlación	Significancia estadística	R <sup>2</sup>	F-Statistic
Temperatura	0.383	$1.240 \times 10^{-14}$	0.3	$3.9 \times 10^{-24}$
Intensidad luz UV	0.375	$5.650 \times 10^{-06}$		
pKa ácido	0.140	$4.587 \times 10^{-4}$		
Longitud de onda	0.006	$3.467 \times 10^{-3}$		
pH	0.040	$4.098 \times 10^{-03}$		
Concentración de catalizador	-0.096	$1.106 \times 10^{-02}$		
pKa básico	-0.041	$1.543 \times 10^{-01}$		
Concentración contaminante	-0.095	$7.340 \times 10^{-01}$		

reducir la variabilidad de los resultados experimentales por este factor. Se detalla en la Tabla 2 la estructura de la base de datos (BD), la cual se conforma de 375 resultados experimentales reportados en 25 artículos de estudio de degradación por fotocátalisis heterogénea. El código necesario fue escrito en lenguaje Python, y ejecutado en el entorno de desarrollo de Google colab, del cual se detallará su implementación en las siguientes secciones.

### 3. Análisis exploratorio de los datos

En esta etapa se calcularon las siguientes medidas estadísticas, para entender la relación de los datos que conforman la BD, mediante las librerías de `pandas` y `statsmodels`, los resultados obtenidos se resumen en la Tabla 3. Se obtuvo la correlación de las variables independientes con respecto a la tasa de degradación, encontrando que estas presentaban intensidades nulas.

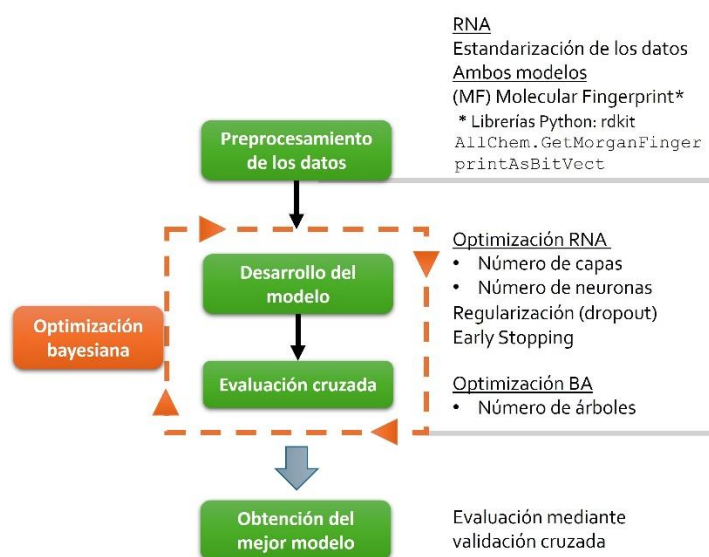
Un algoritmo basado en métodos lineales solo lograría explicar el 30% ( $R^2$ ) de la variable dependiente, por lo que se decide hacer uso de modelos con la capacidad de adaptarse a la relación no lineal entre las variables, como lo son las redes neuronales y los bosques aleatorios.

Se confirma con la prueba de hipótesis F-statistic que existe alguna relación entre las variables independientes y la dependiente, y con base en los resultados de la significancia estadística, se identificaron las variables con impacto en la variable a predecir (aquellas con un valor  $< 0.05$ ). Se omiten para el entrenamiento del modelo las variables pKa básico, potencia de la lámpara y temperatura, en el modelo. pKa básico no cumplió con el criterio anterior (significancia estadística).

Se identificó que la correlación entre las variables intensidad de la luz y potencia de la lámpara es alta (correlación = 0.88), se optó por eliminar la variable potencia de lámpara, ya que se tienen más registros que reportan la intensidad de la luz, y esta es dependiente de la potencia y distancia de la fuente de luz. Se reporta en la literatura que experimentalmente la variación de la temperatura no es un factor que impacte significativamente la eficiencia del proceso, pero se destaca que el proceso de degradación por fotocátalisis solo ocurre entre los 20 y 80 °C [4].

**Tabla 4.** Delimitación de parámetros para el entrenamiento del modelo.

Entrada	Salida
<ul style="list-style-type: none"> <li>Tipo de contaminante (Molecular Fingerprints)</li> <li>Concentración de contaminante                             <ul style="list-style-type: none"> <li>pKa ácido</li> </ul> </li> <li>Concentración de fotocatalizador                             <ul style="list-style-type: none"> <li>pH inicial</li> <li>Intensidad de luz UV</li> </ul> </li> <li>Longitud de onda de la fuente de luz</li> </ul>	Tasa de degradación



**Fig. 1.** Metodología para el desarrollo del modelo

#### 4. Desarrollo de modelo

La metodología implementada se resume en la Figura 1, los parámetros considerados para el entrenamiento del modelo se listan en la Tabla 4.

Para que ambos modelos pudieran reconocer al tipo de contaminante, no como un dato categórico, si no con información interpretable de su estructura química, se obtiene del SMILE (por sus siglas en inglés Simplified Molecular Input Line Entry Specification) el vector binario Molecular Fingerprint de cada uno de los compuestos que conforman la BD.

Para el modelo de RNA fue necesario la estandarización de los datos para un adecuado proceso de entrenamiento, en el modelo de BA no es necesaria la

**Tabla 5.** Resultados de desempeño de ambos modelos.

	RNA	BA
<b>Configuración</b>	<ul style="list-style-type: none"> <li>• Número de capas = 7</li> <li>• Número de neuronas = 50 - 250</li> <li>• Función de activación = ReLu</li> <li>• Optimizador = Adam</li> <li>• Tasa aprendizaje = 0.00094</li> </ul>	<ul style="list-style-type: none"> <li>• Número de árboles = 737</li> </ul>
<b>R<sup>2</sup></b>	0.79	0.81
<b>MSE</b>	0.063	0.052
<b>MAE</b>	0.17	0.15
<b>Tiempo de predicción (s)*</b>	0.1733	0.2023
<b>Tiempo de entrenamiento (s) **</b>	9.532	4.978
*Tiempo promedio de 9 predicciones (ejecutado 6 veces)		
**Tiempo promedio de entrenamiento (Validación cruzada 5 segmentos)		

implementación de esta técnica. Los modelos se implementaron de las librerías `scikit learn` (BA) y `Keras` (RNA).

Los modelos de machine learning generalmente requieren de un proceso de optimización de hiperparámetros para ubicar la configuración con el mejor desempeño de predicción, entre las técnicas de búsqueda más populares se encuentran la búsqueda de cuadrícula, la aleatoria y la bayesiana. Para el modelo de RNA se encontró que no era viable realizar una búsqueda exhaustiva (búsqueda de cuadrícula) para un determinado rango de condiciones, debido a los altos tiempos de procesamiento requeridos, así se decide implementar la optimización bayesiana en ambos modelos, permitiendo encontrar una opción con desempeño destacable.

Se optimiza el número de árboles en BA y el número de capas y neuronas en la RNA (mediante la librería `scikit-optimize BayesSearchCV`), se realizan además pruebas de early stopping y dropout en este modelo. La evaluación se realizó por validación cruzada (5 segmentos), durante el proceso de optimización, y para la evaluación del desempeño de predicción de los modelos, con el fin de reducir al mínimo el sesgo relativo a la división del conjunto de datos.

## 5. Resultados

Se presenta en la Tabla 5 la comparativa del desempeño de los mejores modelos obtenidos. Considerando los resultados, se observa que BA tiene una pequeña ventaja sobre la predicción en comparación con la RNA, sobre los tiempos, se observa que RNA tarda menos en realizar predicciones con nuevos datos, en comparación de BA, respecto al tiempo de entrenamiento, la RNA tarda significativamente más que BA. En la práctica, el tiempo de entrenamiento también impacta en el tiempo de búsqueda de los mejores hiperparámetros, además, la búsqueda en el modelo de RNA es más amplia

que BA (ya que hay un mayor número de combinaciones posibles), se observó que la optimización de RNA tardo significativamente más que BA.

## **6. Conclusiones**

El análisis y desarrollo propuesto de los modelos que presentaron el mejor desempeño permitió obtener una predicción de la tasa de degradación alrededor del 80% con 7 variables, adaptándose al comportamiento no lineal de las variables con respecto a la variable dependiente.

El trabajo se abordó desde 3 perspectivas: la optimización del desempeño del modelo, el contenido de la base de datos y la selección de los parámetros relevantes que intervienen en el proceso.

En la optimización del desempeño del modelo se busca implementar las herramientas necesarias que permitan que el algoritmo se adapte al patrón de comportamiento de los datos. Se debe buscar que el contenido de la base de datos proporcione la información suficiente y de calidad para distinguir en ella el patrón de comportamiento de estudio.

Por último, se deben considerar todas las variables involucradas en el proceso de estudio necesarias para la caracterización de su comportamiento (todas aquellas que impactan en la variable a predecir). Bajo esta perspectiva se han logrado los resultados obtenidos. La predicción obtenida hasta el momento puede presentar mejoras para disminuir el error de predicción si se añaden datos a la BD, si se considerando la existencia de un posible sesgo por la poca presencia de resultados experimentales expuestos a una variación limitada de las variables de experimentación de cada contaminante. Una BD de calidad es de suma importancia para un adecuado entrenamiento de los modelos de aprendizaje automático, esta no solo se limita a que los datos en ella sean fidedignos, sino que exista en ella información suficiente para identificar patrones adecuadamente. Si bien fue un reto encontrar artículos que reportaran la información requerida, por lo que aumentar la BD requiere la inversión de un tiempo considerable, el trabajo tiene el potencial de sentar las bases y consideraciones de metodología para la adquisición de datos experimentales en laboratorio para la generación e implementación de modelos con aplicaciones similares al presentado.

## **Referencias**

1. Rodríguez, D.J., Serrano, H.A., Delgado, A., Nolasco, D., Saltiel, G.: De residuo a recurso: cambiando paradigmas para intervenciones más inteligentes para la gestión de aguas residuales en América Latina y el Caribe. Banco Mundial, Washington, DC (2020). DOI: 10.1596/33436.
2. Chen, D., Cheng, Y., Zhou, N., Chen, P., Wang, Y., Li, K., Huo, S., Cheng, P., Peng, P., Zhang, R., Wang, L., Liu, H., Liu, Y., Ruan, R.: Photocatalytic degradation of organic pollutants using TiO<sub>2</sub>-based photocatalysts: a review. *Journal of Cleaner Production* 268, 121725 (2020). DOI: 10.1016/j.jclepro.2020.121725.
3. Bertagna Silva, D., Buttiglieri, G., Babić, S.: State-of-the-art and current challenges for TiO<sub>2</sub>/UV-LED photocatalytic degradation of emerging organic micropollutants.

*Katia L. Rocha-Aguilar, José I. Hernández-Vega, et al.*

Environmental Science and Pollution Research 28(1), 103–120 (2021). DOI: 10.1007/s11356-020-11125-z.

4. Lee, D.-E., Kim, M.-K., Danish, M., Jo, W.-K.: State-of-the-art review on photocatalysis for efficient wastewater treatment: attractive approach in photocatalyst design and parameters affecting the photocatalytic degradation. *Catalysis Communications* 183, 106764 (2023). DOI: 10.1016/j.catcom.2023.106764.
5. Malviya, A., Jaspal, D.: Artificial intelligence as an upcoming technology in wastewater treatment: a comprehensive review. *Environmental Technology Reviews* 10(1), 177–187 (2021). DOI:10.1080/21622515.2021.1913242.

Electronic edition  
Available online: <http://www.rcs.cic.ipn.mx>



<http://rsc.cic.ipn.mx>



Centro de Investigación  
en Computación