REVIEW

Check for updates

Photocatalytic degradation of drugs and dyes using a maching learning approach

Ganesan Anandhi and M. Iyapparaja [ID] *

The waste management industry uses an increasing number of mathematical prediction models to accurately forecast the behavior of organic pollutants during catalytic degradation. With the increasing quantity of waste generated, these models are critical for reinforcing the efficiency of wastewater treatment strategies. The application of machine-learning techniques in recent years has notably improved predictive models for waste management, which are essential for mitigating the impact of toxic commercial waste on global water supply. Organic contaminants, dyes, pesticides, surfactants, petroleum by-products, and prescription drugs pose risks to human health. Because traditional techniques face challenges in ensuring water quality, modern strategies are vital. Machine learning has emerged as a valuable tool for predicting the photocatalytic degradation of medicinal drugs and dyes, providing a promising avenue for addressing urgent demands in removing organic pollutants from wastewater. This research investigates the synergistic application of photocatalysis and machine learning for pollutant degradation, showcasing a sustainable solution with promising effects on environmental remediation and computational efficiency. This study contributes to green chemistry by providing a clever framework for addressing present-day water pollution challenges and achieving era-driven answers.

1. Introduction

Dye and pharmaceutical wastewater often contains a wide variety of chemicals and compounds. Effluents from the dye industry typically contain trace amounts of dyes,[1] solvents,[2] and heavy metals,[3] whereas effluents from the pharmaceutical industry may contain pharmaceutical residues, organic solvents,[4] and other pollutants (Fig. 1).[5] Without effective management and treatment, these wastes can pose threats to human health and the environment.[6,7] For instance, medication–production effluents containing pharmaceutical residues can have unfavourable effects on aquatic ecosystems and may contribute to the development of antibiotic resistance. Similarly, chemicals and dyes used in the textile and dye industries can negatively impact aquatic life and water quality. When releasing wastewater into the environment or the municipal sewage system, it is often necessary to employ specialized treatment procedures to eliminate or neutralize these toxins.[8,9] To protect human health and the environment, these sectors require effective wastewater management. In the field of waste management, there is a growing need for predictive models that can accurately determine the catalytic degradation performance of organic pollutants.[10–12,117] These models can help design and optimize wastewater treatment processes, ultimately leading to more effective and efficient removal of harmful contaminants from the environment.[13] Machine learning algorithms have emerged as powerful tools in developing predictive models for waste management[14–16] and toxicity prediction.[17] According to a research study by Abdallah *et al.*, several machine-learning

Fig. 1 Various types of wastewater contaminants.

Department of Smart Computing, School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore 632014, Tamil Nadu, India. E-mail: iyapparaja.m@vit.ac.in

methods have been used to create artificial intelligence models connected with garbage control.[18] Artificial neural networks,[19] support vector machines,[20] direct regression examination,[21] decision trees,[22] and genetic algorithms[23] are commonly used techniques to achieve this goal.[24] A novel algorithm has been steadily adopted to better manage waste and encourage environmentally sound expansion because of its demonstrated ability to display complex nonlinear phenomena.[25] Artificial intelligence techniques have also achieved remarkable results in developing forecasts for waste control owing to their capacity to simulate intricate mechanisms.[19,26] Forecasting models for waste control, specifically concerning the catalytic deterioration efficacy of natural contaminants, are crucial for ensuring efficient and sustainable wastewater treatment procedures.[27,28] Artificial intelligence algorithms have increasingly been employed in studies on solid waste control to construct models for predicting diverse waste-related variables. These variables encompass the heat-related behaviour of solid waste, local waste production, and solid waste characteristics such as a greater calorific value.[29–31] Artificial intelligence algorithms can be used to study and comprehend the intricate connections among different factors that affect the deterioration efficacy of natural contaminants.[32–35] Hence, this information can be used to optimize the design and operation of wastewater treatment systems, ultimately leading to the more efficient removal of organic pollutants from the environment.[36–38] Machine learning algorithms play a crucial role in predicting the catalytic degradation performance of organic pollutants in wastewater treatment processes. They enable researchers and engineers to analyze large amounts of data and identify patterns and relationships that may not be easily discernible through traditional analytical methods.[39–45] Using machine learning algorithms, researchers can develop accurate models that can predict the performance of catalysts in degrading organic pollutants.[46–48] These predictive models can then be used to optimize the selection and design of catalysts, leading to more efficient and cost-effective wastewater treatment processes.[49–51] Furthermore, machine learning algorithms in waste management can also assist in decision-making processes for waste treatment and disposal.[52–57] Machine learning algorithms help in identifying the most appropriate treatment technologies and strategies for different types of waste, considering factors such as waste composition, environmental impact, and cost-effectiveness.[25,58–62]

Because machine learning algorithms are suitable for depicting complex nonlinear processes, they are gradually being adopted to better manage waste and facilitate sustainable environmental development.[63–67] These algorithms can process massive datasets and discover previously hidden patterns and discernible relationships through traditional analytical methods. Machine learning algorithms such as artificial neural networks, support vector machines, decision trees, and genetic algorithms have been commonly used to develop predictive models for waste management and wastewater treatment.[68–74] These models are trained using historical data on the performance of catalysts and other relevant parameters such as temperature, pH, and concentration of pollutants.[75–81] Once

trained, these models can accurately predict the catalytic degradation performance of organic pollutants, allowing researchers and engineers to optimize the selection and design of catalysts for efficient wastewater treatment.[82–87] Machine learning algorithms have also been used to examine the hydraulic conditions for efficient flocculation during wastewater treatment.[79,88–91] The integration of machine learning algorithms with wastewater treatment processes offers numerous benefits.[92] It can improve the overall efficiency and effectiveness of the treatment process, leading to higher rates of pollutant degradation and cleaner water. In addition, machine learning algorithms can assist in reducing energy consumption and operational costs by optimizing plant the performance and minimizing resource wastage.[93] Furthermore, machine learning algorithms can provide valuable insights into the behaviour and dynamics of wastewater, allowing for better decision-making in waste management strategies.[94–99] This fusion additionally assists in the improvement of ecological conservation by supporting the recognition and execution of eco-friendly measures in wastewater treatment. Artificial intelligence algorithms can assist in identifying the most efficient triggers and optimizing their efficiency, thereby reducing the environmental influence of organic contaminants.[10,79,100–103] Moreover, these algorithms can assist in predicting the catalytic breakdown effectiveness of organic pollutants, enabling the selection of the most effective treatment techniques.[104,105] In general, artificial intelligence algorithms have been demonstrated to be effective tools for forecasting the catalytic breakdown effectiveness of organic contaminants in wastewater treatment.[106] Machine learning (ML) techniques have become increasingly valuable in investigating the photocatalytic degradation processes of chemicals and chemical compounds. These techniques use computational power to analyze complex data, identify patterns, and make predictions, and help researchers optimize photocatalytic degradation processes. A brief introduction to how machine learning can be applied in this context is given below.

(1) Data processing and feature extraction
(2) Predictive modelling
(3) Optimization and design of experiments
(4) Pattern recognition and mechanism elucidation
(5) Real-time monitoring and control
(6) Data integration and knowledge discovery

This overview sets the context by highlighting the increasing significance of drug and dye degradation processes in the environmental and healthcare sectors. It explains the challenges associated with conventional degradation methods and introduces the potential of ML to enhance the efficiency, accuracy, and predictability of the degradation processes. This review manuscript holds great importance for several reasons. Emerging environmental concerns,[107] optimization of degradation processes,[108–110] reduced experimental costs and time,[111] enhanced predictive capabilities. Numerous studies have explored this domain and offered diverse perspectives. For instance, Ghaedi and Vafaei (2017) predominantly focused on dye adsorption modelling using artificial neural network (ANN) techniques. Interestingly, only a limited array of optimizers, including the Adaptive Neuro-fuzzy Inference System (ANFIS),

support vector machine (SVM), Particle Swarm Optimization algorithm (PSO), and Genetic Algorithm (GA), were employed, but these applications were relatively narrow and lacked in-depth elucidation or graphical representation. Furthermore, the evaluation metrics were limited, encompassing the mean squared error (MSE), R^2 (coefficient of determination), sum of squared errors ARE, X^2, Root Mean Squared Error (RMSE), and Society of Automotive Engineers (SAE), with scant justification for their selection. Future research objectives are also sparse, totaling only three distinct goals.[112] Similarly, Fan *et al.* (2018) primarily concentrated on a specific subset of dye types, with a predominant emphasis on the removal of a wide spectrum of pollutants. Most studies favour ANN, with only a few incorporating Genetic Algorithms (GA). Intriguingly, when it came to degradation modelling, a notable disparity emerged between the Response Surface Methodology (RSM) and Random Forest (RF) model.[113] In 2021, Liu *et al.* introduced a novel approach, presenting a three-layer artificial neural network (ANN) with various architectures for dye analysis.[114] Bhagat *et al.* (2023) highlighted the constraints faced by both ANN and RSM owing to optimization challenges within the degradation process.[115] A comprehensive examination of pollution employing machine learning techniques such as ANN, RF, and Support Vector Machine (SVM) was undertaken by Taoufik *et al.* during the same year. Emphasis was placed on understanding the physicochemical transformations that various pollutants undergo, which is a paramount research priority.[116] In 2023, Bhagat *et al.* contributed to this field by conducting an exhaustive exploration encompassing degradation modelling across all categories of dye types and the corresponding removal methodologies. Their approach also integrated economic analysis to provide a holistic evaluation, addressing the potential pitfalls of misleading data visualization. Consequently, their work served to enhance our understanding of this complex landscape, underscoring the economic implications of data availability and offering a logical interpretation of costs.[13] This paper explores into the fascinating realm of drug and dye molecule degradation, exploring how various machine learning techniques, such as ANN, Levenberg–Marquardt, linear regression, gradient boosting regression, random forest regression, genetic algorithms, Cat Boost, gradient boosting, Hist Gradient Boosting, Extra Trees, XG Boost, decision trees, bagging, light gradient boosting machines, Gaussian processes, and combinations thereof, influence the degradation processes. This study accurately investigated the consequences of employing these diverse machine learning methodologies on the degradation behavior of drug and dye molecules. The overarching objective is to unearth valuable insights into the most efficacious algorithms for predicting and comprehending intricate breakdown mechanisms. By doing so, this research aims to facilitate the development of more streamlined and sustainable drug and dye molecules, thus contributing to the advancement of the field and addressing the environmental and health concerns associated with pharmaceutical and dye waste. It highlights the interdisciplinary nature of the field and promotes collaboration among experts in chemistry, environmental science, and computer science.

2. Photocatalyst

Nanomaterials can significantly alter a product's processing, design, and characteristics. Capabilities, such as optical, magnetic, electric, antibacterial, and mechanical capabilities, are all significantly affected when a material is reduced in size from macro to nanoscale. This is mainly triggered by an increase in the surface-area-to-volume ratio and surface reactivity, which have an impact on the chemical and physical characteristics of the material. Materials with at least one dimension between one and 100 nm are referred to as nanomaterials. These materials have made significant contributions to electrical and information technologies by enabling the development of powerful, easy-to-use, and energy-efficient products. These abrupt alterations in material characteristics at the nanoscale have enormous potential for applications in a variety of sectors, including photocatalytic degradation and storage.[128,155]

Photocatalysts are semiconductors that catalyze reactions when exposed to light. Unlike other energy sources, photocatalysts have become the standard for green catalysts in recent years because they are non-hazardous. Photocatalysis dates back to the early 20th century, when studies explored the photochemical reactions of titanium dioxide (TiO_2) and other semiconductor materials. In the 1970s and 1980s, researchers investigated the potential of photocatalysis for environmental remediation, including water purification. The pioneering work by Fujishima and Honda in 1972 on the photoelectrochemical properties of TiO_2 laid the foundation for modern photocatalysis. In semiconductor materials, an electron–hole pair is created when exposed to light. One of the main factors influencing the physical properties of semiconductors is their energy bandgap. The difference in energy between the valence band (HOMO) and conduction band (LUMO) is represented by the energy band gap (E_g).[156] These properties include band gap values, recombination rates, and the availability of oxidizing agents. However, the photocatalyst preparation parameters and pollutant properties must be considered for an excellent photocatalytic reaction.[157]

Photocatalytic degradation is a promising technique for removing organic pollutants from water. This involves the use of photocatalysts to accelerate the degradation of organic compounds under light irradiation. This process is particularly important for the degradation of dyes and drugs, which are common pollutants in industrial effluents and wastewater. Despite its potential, photocatalytic degradation faces several challenges, including the limited efficiency of photocatalysts under visible light, the need for high-intensity UV-active catalysts, and the potential for the formation of toxic byproducts during degradation. Additionally, the scalability and practical implementation of photocatalytic systems for large-scale wastewater treatment remain areas of concern. Fine-tuning parameters such as catalyst loading, pH, temperature, and light intensity are crucial for improving the degradation efficiency and reaction kinetics. Combining photocatalysis with other treatment methods such as ozonation or sonolysis has shown synergistic effects, leading to enhanced degradation

rates and broader substrate specificity. Advances in kinetic modelling and mechanistic studies have provided insights into degradation pathways and reaction kinetics, aiding in the optimization and scale-up of photocatalytic processes.

Recent research efforts have focused on addressing these challenges and improving the efficiency and applicability of the photocatalytic degradation processes. Strategies such as doping with metals or nonmetals (i) metal and non-metal doping, (ii) co-doping, (iii) composites, (iv) substitution, (v) sensitization, and (vi) various other methods, coupling with other materials (such as supporting materials), and optimizing catalyst morphology and surface properties have been explored to enhance the photocatalytic activity and extend the spectral response range towards visible light.[158] Over the years, researchers have explored various semiconductor materials, nanostructures, and dopants to enhance the photocatalytic activity and efficiency. Photocatalytic degradation has applications in various fields, including wastewater treatment, environmental remediation, and wastewater treatment in the pharmaceutical industry. The potential for the removal of emerging contaminants, such as pharmaceuticals and personal care products, from wastewater has garnered significant attention in recent years. Some of the photocatalyst such as Pt/TiO_2, CuO/TiO_2, CuO–MMT, Ni/ZnO, ZnO, Ag–ZnO, ZnO–rGO *etc.* are used for pollutant degradation.

Descriptor selection in ML modelling is important for the photocatalytic degradation of drugs and dyes to capture the relevant factors affecting the process. The examples are as follows.

(a) Molecular structure theorists

Molecular descriptors, such as chemical functional groups, can be used to capture the structural characteristics of drugs or chemical compounds. For example, descriptors representing the presence of specific functional groups (*e.g.*, aromatic rings or functional moieties) in a chemical photocatalyst can influence the rate of degradation.

(b) Behavioral state descriptors

Reaction conditions, such as temperature, pH, and catalyst concentration, play an important role. These descriptions of the degradation process help model the environmental effects on the photocatalytic process and determine the optimal conditions for optimum performance.

(c) Electronic and optical properties

It is important to include descriptors of the electronic and optical properties, such as the energy level and absorption spectrum. For example, in dye photocatalysts, descriptors representing the electronic configuration and absorption of a substance can help predict its degradation under specific photocatalytic conditions.

The careful selection of annotations ensures that ML models capture relevant dye and drug components and their interactions with photocatalysts, giving the models a good ability to predict the photocatalytic degradation processes.

2.1 Kinetics of degradation

Based on the change in the concentration of the pollutant (dye or drug), the kinetics of the photodegradation reactions were analyzed by monitoring the characteristic absorbance peak at various illumination time intervals.

$$\text{Degradation efficiency } (\%) = (C_0/C_0)100$$

where C_0 and C are the solution concentrations at $t = 0$ and after some illumination time intervals, respectively.[159]

2.2 Factors influence photocatalysis

2.2.1 Effect of pH. pH is an important factor for photocatalytic dye decolorization because it has a variety of effects on dye reaction rates. Since the pH of a given solution determines the catalyst surface charge, pH can have an impact on dye adsorption on semiconductor surfaces.[160]

2.2.2 Effect of dye concentration. The adsorption of dyes on the surface of a photocatalyst is necessary for the photocatalytic process to occur. The amount of dye adsorbed on the photocatalyst surface, not the amount in the bulk of the solution, is the factor that drives photocatalysis during photodegradation. The initial dye concentration in each photocatalytic process is an important consideration for dye adsorption. By raising the dye concentration, the percentage degradation falls while maintaining a constant amount of photocatalyst.[161]

2.2.3 Effect of catalyst concentration. Another factor influencing dye degradation is photocatalyst concentration. The rate at which the dye was photodegraded increased with the concentration of the photocatalyst. With an increase in catalyst concentration, the photocatalyst's surface has a greater number of active sites.[162]

2.2.4 Effect of light irradiation. The amount of light and duration of irradiation affected the decolorization of the dye. When photocatalysts absorb light energy equivalent to or greater than the band gap energy, electrons are transferred from the VB to the CB, leaving the VB with holes. Photocatalytic degradation occurred at a faster rate when the radiation intensity increased.

2.2.5 Effect of temperature. As the temperature increases, the rate at which organic molecules photodegrade becomes more efficient. When the solution forms bubbles owing to an increase in temperature, free radicals are produced. Furthermore, the degradation reaction that overcomes electron–hole recombination is aided by the temperature increase.[163]

2.2.6 Effect of nature of the catalyst. Different photocatalysts photodegrade in different ways depending on their lattice mismatch and BET surface variations. Impurities on the photocatalyst surface also affect the adsorption behavior of the effluent as well as the lifetime and recombination rate of the electron–hole pairs. A large surface area can play a significant role in photodegradation. However, depending on the particle size, the electron–hole recombination mechanism can change. Several semiconductor nanocatalysts have been demonstrated to act as photocatalysts for the treatment of wastewater and pharmaceutical pollutants. The influence of photocatalyst efficiency on the operating parameters was studied. Investigating the characteristics of the sample to be degraded is essential because the impact of different parameters has occasionally generated controversy. Because the reaction exceeds pseudo-

first-order kinetics, the rate decreases with illumination time. In addition, there can be a conflict between the reactants and the intermediate products for decolorization.[164]

3. Analysing the performance of machine learning algorithms

Artificial intelligence algorithms play a vital role in forecasting the catalytic disintegration performance of organic contaminants during sewage treatment. These algorithms analyze and interpret data on variables, such as heat, acidity, and contaminant quantity, to formulate models that can precisely predict the efficacy of catalytic disintegration. The combination of artificial intelligence algorithms and sewage treatment processes makes it possible to simulate sophisticated nonlinear linkages, which conventional models may have difficulty depicting because of the complexity of the relationships involved.[118,119] These algorithms can unearth concealed patterns and connections within the data, leading to an improved comprehension of the variables that impact catalytic disintegration performance. By assimilating artificial intelligence algorithms into sewage treatment procedures, researchers and engineers can make more knowledgeable choices concerning catalyst selection, process enhancement, and treatment approaches.[10,120] AI algorithms are skilled at managing extensive and intricate data collections, enabling the examination of a broad spectrum of factors that could potentially influence catalytic disintegration. This capability allows researchers to simultaneously consider multiple variables, resulting in more precise forecasts. Moreover, AI algorithms can simulate nonlinear connections without the requirement for explicit mathematical or chemical formulas.[15,121] Machine learning algorithms have become increasingly valuable tools in environmental science, particularly for the elimination of organic pollutants through catalysis. These algorithms can analyze large datasets and identify patterns and relationships that may not be immediately apparent to researchers.[122,123] Researchers can make efficient predictions regarding the catalytic reduction efficiency of eco-toxic chemicals utilizing machine learning models.[124] For example, Dong *et al.* formulated a machine learning framework with the objective of forecasting the degradation efficiency of organic pollutants through the utilization of carbon-based catalysts. The model considers multiple variables, including the surface area of the catalyst,[125,126] distribution of pore sizes, and functional groups present on the surface. In addition, to anticipate the photochemical oxidative breakdown of phenolic pollutants, Dondapati *et al.* (2020) used a supervised machine learning approach.[127] These studies provide evidence of the potential of machine learning in the prediction of catalytic degradation performance and the identification of highly effective catalysts.[129] However, despite these advancements, there are still challenges that need to be addressed when applying machine learning models to the field of catalytic degradation. One of the challenges encountered in this study pertains to the absence of a correlation relationship and the relative significance of variables.[130] Despite the accurate predictive capabilities of machine learning models in assessing catalytic degradation performance, there remains a lack of comprehensive

understanding about the fundamental connection between catalytic fingerprinting and waste-to-energy conversion.[131] Furthermore, it has been observed that the fingerprint features of catalysts, such as the atomic number, chemical formula, and electronic structure features, may have an influence on the decontamination efficiency, the yield, and the nature of reactive oxygen species.[1,132] However, the task of statistically establishing the inherent correlation involving the characteristics of catalysts' fingerprints and the decomposition of pollutants continues to pose a significant difficulty.[133] In order to tackle these issues, scholars are employing simulations and experiments to quantitatively investigate the inherent correlation between the fingerprint attributes of catalysts and the degradation of pollutants.[134] They are also incorporating advanced techniques such as feature engineering and dimensionality reduction to find set patterns and relationships within the data.[135] By acquiring a more profound understanding of the connection involving the chemical, catalysts and the decomposition of pollutants, it becomes feasible to boost the efficiency of machine learning patterns.[136] This enhancement empowers the accurate anticipation of catalytic efficiency and simplifies the choice of exceptionally efficient catalysts purposed at breaking down organic toxins.[1] Machine learning algorithms have shown promise that this method can predict and select effective catalysts for degrading organic contaminants. These algorithms have been successfully used to predict the catalytic rate constant and degradation efficiency of pharmaceutical compounds using metal–organic frameworks.[148] Furthermore, machine-learning models can consider multiple variables simultaneously, allowing for a more comprehensive analysis of catalyst performance. The utilization of machine-learning algorithms to forecast the catalytic deterioration capability of organic pollutants provides various benefits. Firstly, machine learning algorithms can handle large and complex datasets, allowing for the inclusion of multiple variables that may impact catalytic performance.[137] This ability to consider multiple variables simultaneously enhances the predictive power of the model, leading to more accurate and reliable results. Furthermore, machine learning models can also consider nonlinear connections between the fingerprint features of catalysts and pollutant deterioration, which might not be captured by conventional statistical models. Furthermore, machine learning algorithms can continuously learn and improve new data, making them adaptable to changing conditions and allowing for the development of more robust models over time.

To inspect the effectiveness of machine learning algorithms in forecasting the chemical breakdown efficiency of natural pollutants, scientists have utilized different methods.[138,139] These methods incorporate guided machine learning, where the algorithm is trained on a categorized dataset to grasp the connection between the input factors and the chemical breakdown efficiency, and unsupervised machine learning, where the algorithm distinguishes trends and connections in the data without any previous understanding or categorizations.[153,154]

In the field of computer vision, assessment metrics play an essential role in assessing the overall performance of the gadget models. Commonly used metrics include accuracy, precision, recollection, and $F1$ rating, which can be especially applicable to obligations that include photograph type and object detection.

These metrics provide insights into the potential of the model for efficiently classifying and discovering items inside a photograph. Additionally, Intersection over Union (IoU) is frequently used to measure the overlap between predicted and floor reality bounding bins, imparting a more nuanced evaluation of the localization accuracy. In natural language processing (NLP), research is regularly based on metrics such as accuracy, precision, recall, and $F1$ ratings for tasks such as sentiment evaluation or named entity popularity. Specialized metrics such as BLEU and ROUGE are employed to evaluate the first rate of generated textual content in obligations, which includes device translation and summarization. These metrics serve as critical equipment in quantifying the effectiveness of device-mastering models throughout diverse research domains, facilitating the comprehensive knowledge of their competencies and obstacles.

For example, one of the authors developed a machine learning model to predict personal exposure to benzene.[133] They compared the performance of their machine learning model with a linear regression approach and found that the machine learning model showed higher variability in performance, accurately classifying personal exposure levels to benzene. These studies demonstrated the potential of machine learning methods for accurately predicting the catalytic degradation performance of organic contaminants. By leveraging the power of machine learning, researchers can gain valuable insights into the factors that influence catalytic degradation and predict the performance of organic pollutants under different conditions. Furthermore, it was found that using machine learning algorithms to estimate environmental pollutants yields significant improvements in accuracy. For instance, the use of random forest approach has proven effective in estimating fine particulate matter ambient ozone levels and groundwater nitrate pollution.[17]

AI methods have enabled the quantitative identification of the intrinsic link between catalyst fingerprint properties and pollutant degradation, which was previously difficult to determine.[133] This breakthrough in utilizing machine learning algorithms for catalyst screening and selection significantly improved our understanding of the factors that impact decontamination performance, together with the quantity and variety of reactive oxygen species produced during the degradation process. Estimating the catalytic degradation performance of organic contaminants is critical for creating efficient environmental remediation solutions.

Machine-learning models can also enhance the development of more efficient and sustainable catalytic processes by providing insights into the underlying mechanisms and key parameters that drive pollutant degradation. Overall, machine-learning algorithms have the potential to revolutionize the field of catalytic degradation of organic pollutants.

4. Applications of machine learning in pollution control

Machine learning algorithms have a wide range of applications in pollution control[140,141] air quality epidemiology[142] and environmental remediation.

They can be used as follows.

(1) The effectiveness of different catalysts in degrading organic pollutants was predicted based on their fingerprint characteristics.

(2) Optimize the reaction conditions and process parameters to maximize the pollutant degradation efficiency.

(3) Identify the most influential factors that contribute to the degradation of organic pollutants, allowing for targeted improvements in catalyst design and performance.[143]

(4) Assess the environmental impact of different pollutants and predict their behavior under various scenarios.

(5) Design-tailored treatment plans for specific pollutants based on their chemical properties and characteristics.

(6) Develop real-time monitoring systems to detect and track pollutant levels, allowing prompt response and mitigation measures. By harnessing the power of machine-learning algorithms, researchers can expedite the discovery and optimization of potent catalysts for pollutant degradation. This can lead to significant advancements in pollution control and environmental remediation.

Catalytic conversion reactions require precise control and optimization of the reaction conditions, which can be time consuming and challenging for chemists. Artificial intelligence algorithms can analyze vast amounts of data, identify the most efficient reaction parameters, save time, and improve the overall success rate of these reactions. AI can also predict the outcomes of different catalytic conversions, allowing researchers to make informed decisions and accelerate the development of new drugs, dyes, and wastewater treatment methods.

4.1 Drug degradation

Drugs are degraded by a process known as photocatalytic drug degradation when exposed to light, especially when a photocatalyst is present. A substance known as a photocatalyst is one that speeds up a photochemical reaction by collecting light energy and transferring it to drug molecules, which causes them to be chemically altered. Drugs that are light-sensitive are particularly susceptible to this type of degradation, which can alter their composition, potency, and overall efficiency.

This section discusses the influence of the physicochemical properties of the drug on its degradation, which are heavily influenced by their structural and physicochemical features and can be analyzed using ML methods. The degradation behavior of new pharmaceuticals can be predicted using ML models trained on a dataset of known drug degradation designs. To further improve the efficacy of drug degradation, ML algorithms can be utilized to optimize the photocatalytic reaction parameters, including light intensity and catalyst concentration.

Gordanshekan *et al.*, in 2022, proposed the use of two photocatalytic systems to reduce cefixime, $Bi_2WO_6@TiO_2$ and g-$C_3N_4@Bi_2WO_6$. Adsorption isotherms and photocatalytic breakdown kinetics were studied using UV irradiation. Toxicity estimation, degradation pathway analysis, and Artificial Neural Network (ANN) analyses were performed in this study. The degradation of cefixime and its byproducts was complete after 180 min of the reaction. After 135 min of reaction, LC-MS and

TOC showed that the majority of harmful products were no longer present, and T.E.S.T. verified that safer chemicals had been isolated. The number of hidden neurons along each ANN path also varied from 1 to 24 before the architecture with the smallest MSE was selected. The photocatalytic degradation of cefixime was predicted using an optimized structure, which was created by optimizing the number of neurons. The artificial neural network model has developed reliability in simulating the photocatalytic degradation of cefixime, as demonstrated by the high accuracy of ANN prediction results with experimental data. The excellent results achieved by ANN suggest that it may be possible to rely on the prediction of any photocatalyst, whether it is a pure form or a binary mixture, and at different component weight ratios. The significance of the variables is shown in Garson's equation. According to the experimental findings, the C/L ratio and reaction time are the most important factors. The fact that the significance values for each imported variable are so close together suggests that none of them can be safely removed.[144]

Hosseini *et al.* 2022 reported that Cu with Al/Layered Double Hydroxide with Graphitic Carbon Nitride (LDH@g-CN) degraded tetracycline when exposed to light. Impressive technical, environmental, and economic success was achieved: 96% degradation in 90 min with no secondary pollution. They used the response surface methodology (RSM) and ANN to achieve their goals. Based on the results of the RSM analysis of variance (ANOVA), researchers used Central Composite Design (CCD) methods. The information in the regression coefficients shows the importance of the variables being examined and how they might interact with each other. The F- and P-values of 71.11 and 0.0001, and 5.23 and 0.0897, respectively, showed the importance of the suggested model for simulating data points inside the test domain. The optimal match between the observed data and those used to predict the results of an experiment is indicated by the R^2 and adjusted R^2 values of 0.96 and 0.91, respectively. According to the results of one-way analysis of variance (ANOVA), the following factors had the greatest impact on tetracycline decay by the catalyst: preliminary tetracycline concentration, time of exposure, and photocatalyst dose. The efficiency increased dramatically from the 8.0–30.0 ppm

tetracycline range of photocatalyst dosages. This is because the increased concentration of nanoparticles improved light absorption while decreasing light dispersion. In addition, the performance was poor across the entire spectrum of tetracycline dosing. These are very important and vital in ANN modelling; without them, the model will not work properly with the data. Root-Mean-Squared Error (RMSE) values have been found to be the lowest when there were 8 neurons in the hidden layer, as determined by an iterative procedure that simultaneously determined the number of neurons and training functions in both the training and testing data sets. Consequently, the AVM removal process was modelled using the feed-forward Levenberg–Marquardt (LM) algorithm, which was chosen after further optimizing the number of neurons using the tansig and logsig output stages. When using eight neurons, trainlm as the training function, and tansig as the transfer function, the RMSE was minimized across all the data series used for training, testing, and validation. After reaching a peak in the $4:8:1$ configuration, the R^2 values began to decline. The rate at which R^2 values shift varies between the datasets used for testing and those used for validation. We decided to use a network with a generalized model for training, validation, and testing because of its low RMSE and R^2 values, which were close to 1 and 0.00102,[145] respectively. Yu *et al.* in 2023 published machine-learning-based catalyst screening and layered double hydroxide-based catalysts for norfloxacin breakdown. A unique simulation-experiment-prediction framework was used to determine how ROS changes at the atomic level and to find stable catalysts for hetero-EF reactions. The reaction time was reduced by almost 2 h owing to the Co_3Fe_2-LDH's outstanding catalytic efficiency, and its $1O_2$ production improved by more than 20% during norfloxacin degradation. They investigated which catalyst conversions for ecological restoration were the most effective. An ML system uses models for future outcomes using one of three popular machine-learning techniques: linear regression (LR), gradient boosting regression (GBR), and random forest regression (RFR), as shown in Fig. 2. Monte Carlo cross-confirmation, on the other hand, was performed to choose a suitable ML technique for this study in light of the small amount of data collected and previous research in this
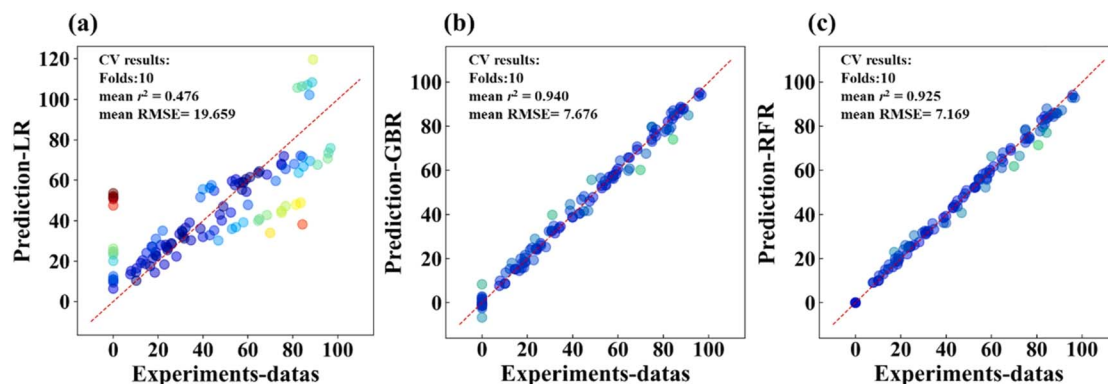

Fig. 2 Performance evaluation of alternative models trained using diverse machine-learning techniques: (a) linear regression (LR), (b) gradient boosting regression (GBR), and (c) random forest regression (RFR). Copyright (2023) with permission from Elsevier.

area. With a cross-validation fold of 10, the data were randomly split into 10 subsets, with the goal of predicting the characteristics from the remaining set. Among the three methods, the GBR algorithm's $R^2 = 0.940$ prediction accuracy of the GBR algorithm was the most impressive. Compared to our model, other approaches failed to accurately predict the degradation time of the drug. The RFR model exhibits a remarkable coefficient of agreement ($R^2 = 0.925$). The prediction accuracy of the LR model was subpar with $R^2 = 0.476$. These models have shown great potential in accurately predicting the degradation time of drugs. Additionally, the use of fingerprint traits of catalysts allows for a cost-effective and efficient approach to GBR forecasting, making it a valuable tool in studying various catalysts for drug removal.[132]

According to a study by Shang *et al.* in 2022, antibiotic pollutants in wastewater must be degraded effectively. The advantages of the photocatalytic breakdown of organic contaminants includes their tolerance to the ecosystem, safety, and completeness. In this study, composite photocatalysts (TS-1/C_3N_4) composed of carbon nitride and titanium silicon molecular sieves were used to break down ofloxacin in wastewater through photocatalysis. Using experimental data and an ANN trained with a genetic algorithm (GA), the reaction parameters were optimized to achieve the best possible results. Under the best testing conditions (1.55 g L^{-1} catalyst, 58.60% per TS loading, and 49.38 mW cm^{-2} luminous power density), the maximum removal efficiency (RE) was measured at 82.92%. Moreover, both experimental and modelling efforts utilizing an ANN model were made to examine the impact of wastewater elements on RE. Experiments showed that when wastewater constituents were present, reductions were observed in both the RE and the amount of ofloxacin adsorbed on the photocatalyst, most likely as a result of the competitive adsorption of the wastewater constituents, interaction combined with the effects of light-blocking and reactive species. In this case, the capacity of the model to forecast the synergistic interplay of wastewater components was demonstrated by an absolute relative deviation (ARD%) of 6.88%, 1.04%, and 1.77% owing to the complementary action of metal ions, anions, and cations, respectively. This research has the potential to yield useful information and methods for the photocatalytic treatment of sewage contaminated with antibiotics.[146]

4.2 Dye degradation

The use of dyes, a prevalent pigment in numerous industries, has posed notable challenges concerning both human well-being and the safeguarding of ecosystems. Hence, the degradation of these pigments is vital for the preservation of life. Moreover, machine learning (ML) models are currently experiencing remarkable progress in their capability to precisely forecast the course of dye photodegradation. In this investigation, a diverse set of machine learning models was formulated to effectively predict the photodegradation of dyes on nanocomposites. Jaffari *et al.* in 2023 published an ML technique to predict how well photocatalysts made of bismuth ferrite ($BiFeO_3$) will work in eliminating malachite green (MG).

Multiple machines learning models shown in Fig. 3, including Cat Boost, gradient boosting, Hist Gradient Boosting, Extra Trees, XG Boost, decision tree, bagging, light gradient boosting machine (LGBM), Gaussian process, the combination of ANN and lightest (a variant of L-S-T-M) are studied in this article to determine their relative strengths and weaknesses. These models were used to determine the efficacy of photocatalytic decay in reducing malachite green concentrations in wastewater using a variety of NM–$BiFeO_3$ composites. Under different settings, 1200 data sets were collected to form a comprehensive database. Several photocatalyst inputs included the catalyst type, reaction speed, illuminance, amount, catalyst dosage, pH condition, humic acid level, anions, surface energy, and average pore size. The rate of MG dye breakdown was chosen as the dependent variable. According to the results of the performance analysis, the cat-boost model is above all others because it has the greatest testing, the correlation coefficient is 0.99, the mean absolute error is 0.64, and the root-mean-squared error is 1.34. Cat Boost can accurately forecast over 96% of the data points, with an absolute relative error of less than 1%, as shown by the cumulative frequency plot. Cat Boost demonstrated that the conditions are more critical than the material used to initiate the photocatalytic reaction. Based on the simulations, the optimal process conditions were as follows: 105 W of light, 1.5 g of catalyst per liter of solution, 5 mg of MG dye at the beginning, and a pH of 7. Furthermore, time was identified as the most important input by feature importance analysis.[147] A carbon-based metalorganic framework (AC-MIL-88B) was investigated by Mahmoodi *et al.* in 2019 and the least-squares support vector machine (SVM) technique was employed to study its potential for the photocatalytic decay of Reactive Red 198 dye (RR198). The photocatalytic impact of AC and MIL-88B (Fe) on the decay of RR198 is superior to that of AC and MIL-88B (Fe). The AC/MIL-88B (Fe) composite catalyzed the photocatalytic degradation of RR198 at 99.9% efficiency using a catalyst amount 0.08 g L^{-1}, pH about 3, and dye concentration 50 mg L^{-1}. The RR198 treatment was performed in agreement with the second kinetic theory. The rate constants (in L mg^{-1} min^{-1}) for 0.2, 0.3, 0.4, and 0.5 g of AC/MIL-88B were 0.0032, 0.0044, 0.0084, and 0.0148, respectively (Fe). As a suggestion, AC/MIL-88B (Fe) is a promising material for the photocatalytic decolorization of RR198 owing to its high degradation ability. The sophisticated LSSVM model might replicate the deterioration in RR198 observed in the experiments. These results show that the model obtained a moderate RMSE, AARD, and correlation ($R^2 = 0.948$), indicating that its predictions were in good agreement with the experimental results (Fig. 4). These findings demonstrate that the LSSVM model captures the dye breakdown *via* photocatalysis using AC/MIL-88B (Fe). Thus, this model shows promise as a useful instrument for forecasting and optimizing the decolorization process in future research.[148] By conducting multiple experimental runs of photocatalytic activity, Ayodele *et al.* in 2021 demonstrated that the collected data can be used with data-driven machine-learning modelling strategies such as ANN (Fig. 5). From the viewpoint of a Levenberg–Marquardt-trained ANN, chloramphenicol, phenol, azo dye, gas styrene, and methylene blue were investigated for their photocatalytic
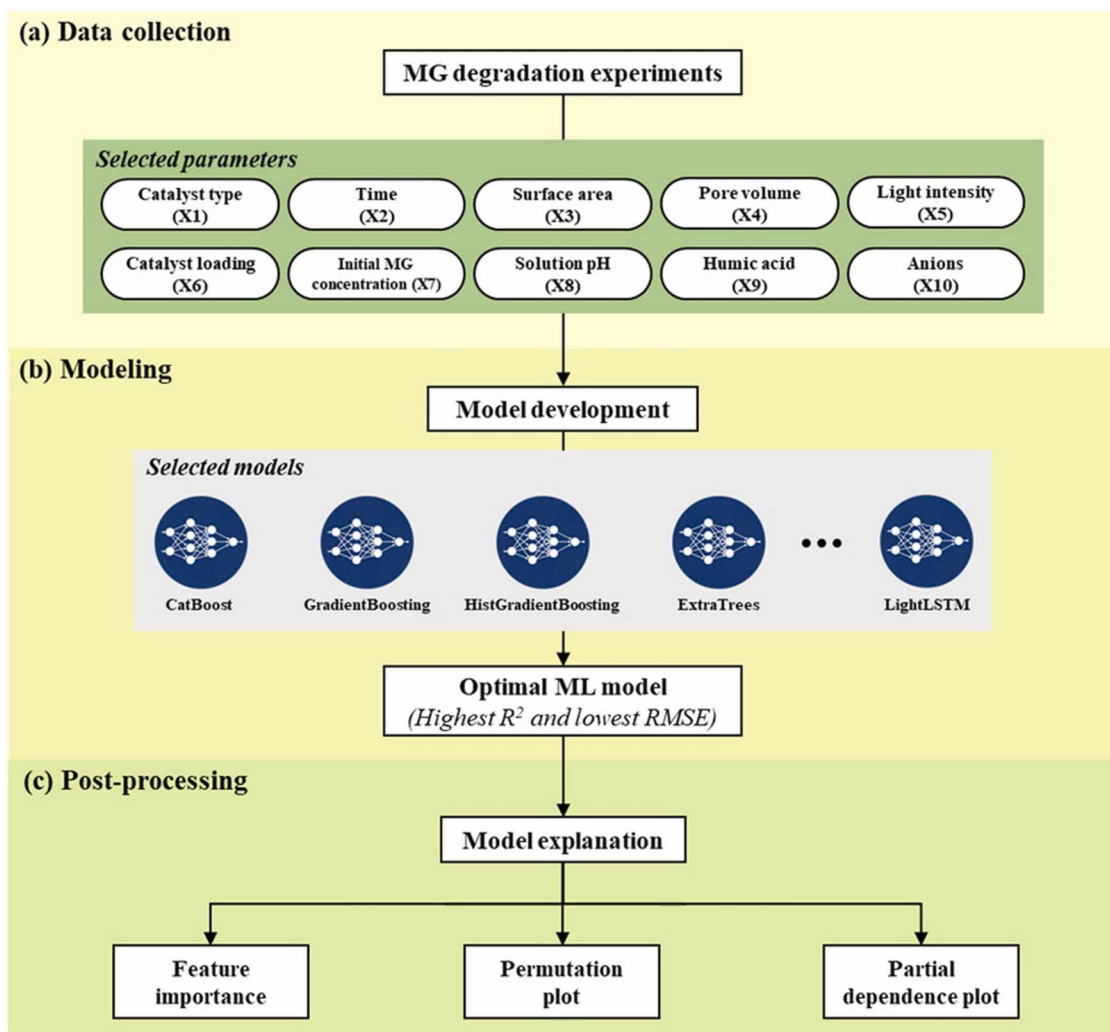
Fig. 3 Illustrative representation of the procedural framework: (a) gathering data, (b) constructing models, and (c) subsequent processing. Copyright (2023) with permission from Elsevier.

breakdown. In each photocatalysis process, the hidden neurons in the 20 distinct neural network topologies were optimized. To show how chloramphenicol, phenol, azo dye, gaseous styrene, and methylene blue break down in sunlight, the optimal ANN configurations obtained were 320-1, 35-1, 32-1, 417-1, 46-1, and 310-1. At the 95% confidence level, the optimized ANN designs had a high R^2 for forecasting the degradation of organic contaminants, and their mean absolute errors were low. Using a modified version of the Garson algorithm, they conducted a sensitivity analysis, which showed that all the process parameters had a substantial impact on the breakdown of organic pollutants by light. Significant influences on photo-degradation were identified as starting methylene blue concentration and variables such as photocatalyst ratio, phenol amount, pH, and hydrothermal temperature. It is important to consider the nonlinear relationship between the process parameters and the resulting degradation rate when designing a photocatalytic reactor to effectively degrade the extent to which these organic pollutants influence the photocatalytic

degradation process. The ANN algorithm can be used in pho-tocatalytic degradation to make crucial decisions.[143]

Sathiskumar *et al.* in 2023 analyzed the different types of dyes and drugs used in the literature. The following nitro-phenols (4NP) were studied: 2,4,6-trinitrophenol, dinitrophenol (DNP), methylene blue, methyl orange, and 2,4,6-trini-trophenol. Several ML algorithms have been used to predict catalytic characteristics based on trial time, including linear regression (LR), support vector machines (SVM), gradient-boosted machines (GBM), random forest (RF), and XGB tree (XGB). They discovered that the NP and DNP were the best XGB algorithm parameters. The SVM model exhibited the lowest RMSE, MAE, and MAPE metrics for the validation dataset. In both the test and training sets, the SVM outperformed the rival models. The GBM and XGB tree models outperformed the LR and RF models on the testing dataset with lower RMSE, MAE, and MAPE values. However, in the training dataset, the RMSE, MAE, and MAPE values for the GBM and XGB tree models were relatively high. The PdO–NiO bimetallic catalyst decreased the
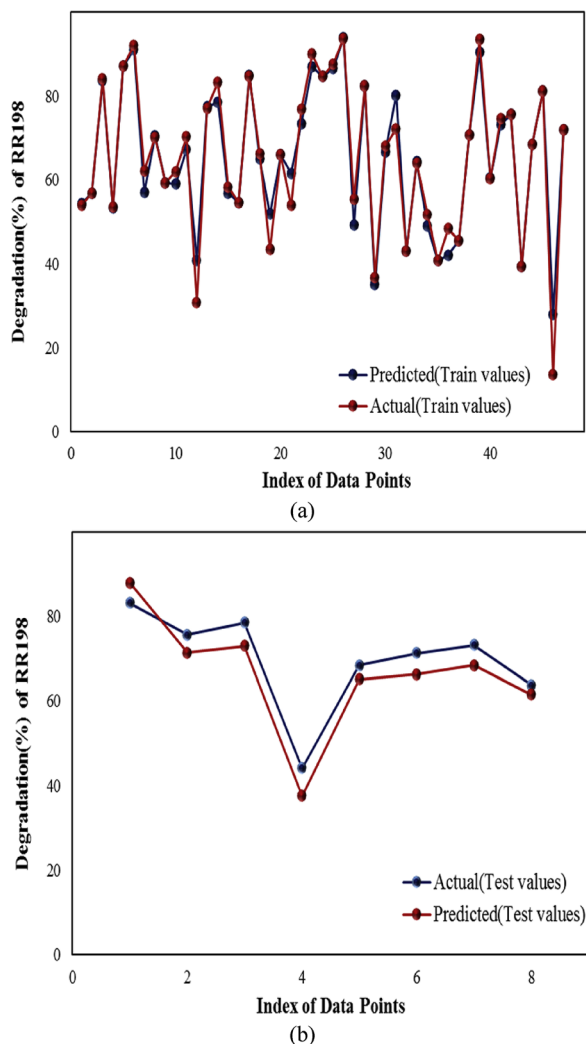
Fig. 4 (a and b) Shows the real data compared with the predicted data in the COA-LSSVM model across the data index. Copyright (2019) with permission from Elsevier.

azo compound mixture by 98% in 8 min. As a result of their experiments, PdO–NiO proved to be an excellent catalyst.[123]

A novel ML-based photocatalyst has been described by Zhai *et al.* in 2023, which effectively degraded the RHB dye using Bi_2WO_6/MIL-53 (Al). The catalytic activity of the BWO/MIL composite materials can be measured primarily by observing the degradation of RhB (Fig. 6). To hasten the development of BWO/MIL that exhibits the desired performance, a nano-photocatalyst module was developed using machine learning. The support vector regression (SVR) forward feature selection approach was used previously, and four important features pertaining to the simulated conditions of BWO/MIL were extracted from the RhB dataset. Second, they identified salient properties and best-fit hyperparameters for a support vector regression (SVR) model to predict the RhB of BWO/MIL. Using an outside test and a technique called leave-one-out cross-validation (LOOCV) yielded R^2 values of 0.823 and 0.884, respectively, for the predicted and experimental RhB. To investigate the synthesis space, they used inverse projection, a prediction model, and virtual screening for

BWO/MIL nanocomposites with increased RhB content, which brought us to our third point. The BWO/MIL composites enhanced the toxic removal during the degradation process and improved the photocatalytic performance. Experimentally synthesized efforts to create highly efficient photocatalysts can be directed by the projected RhB values.[31]

Rodrigues *et al.* (2020)[149] demonstrated the mathematical modelling of the photocatalytic reaction of reactive blue 19 (RB 19) and reactive blue 21 (RB 21) dyes in industrial effluents using ZnO. The photodegradation of the colors in the effluent was maintained by the ZnO catalyst. The ZnO catalyst had the following characteristics: density of 5550 kg m^{-3}, mean particle diameter of 1.19107 m, surface area of 16.830 m^2 g^{-1}, and porosity 0.1 cm^3 g^{-1}. ZnO nanoparticles were found to significantly accelerate the photodegradation of RB 19 and RB 21 over a 6 h period, with efficiencies of 100% and 91%, respectively. The mass balance conservation law can be used to numerically model the reactor and the recycling lines of a photocatalytic reactor. A first-order photocatalytic reaction of pollutant concentration and UV light attenuation is proposed, along with spherical and uniformly sized particles, constant temperature, Langmuir isotherm between the solution and catalyst, constant void fraction in the reactor, and constant reaction time.

The model had a mean absolute error of less than 1.5% and as high as approximately 7% when compared to the experimental deterioration data. The model showed that as the fraction of photocatalytic degradation increased, the concentration of pollutants decreased, and the size of the catalyst increased. This shows that the efficiency of the photocatalytic process greatly depends on the Langmuir isotherm equilibrium and the quantity of free space in the reactor. The results also show the possibility of enhancing pollutant breakdown by adjusting both pollutant concentration and catalyst size.[149]

Phyto-mediated nanoparticles were synthesized and used to produce hydrogen on $NaBH_4$, biological properties, and photocatalytic decay of dyes; the creation of a machine learning model was published in 2022 by Lin *et al.*[141] For the purpose of degradation, they produced an N@Pt–Ag BNP catalyst. The opposition-based principle was used to modify the conventional normal distribution method (OBL). The maximum number of hidden layer neurons, mutation method, and basic variants of DE are all set at the outset of the process. Each generation produces a new individual that encodes information about the neural network of the next generation, and each network (evaluated *via* decoding) calculates DE's fitness function of the DE. This function is then used by the algorithm to select the best candidate for the subsequent generation.

Fifty simulations were run with the following parameters adjusted to find the optimal model for the system under consideration: generations = 500, individuals' population = 40, the most possible number of hidden layers is two; the most possible number of neurons in the first hidden layer is 20, and the most possible number of neurons in the second hidden layer is ten. Data from each of the 50 runs, the optimal ANN contained only one layer, consisting of 15 neurons. The average absolute error was determined as the best model. The percentage is 0.651% during training and 0.579% during
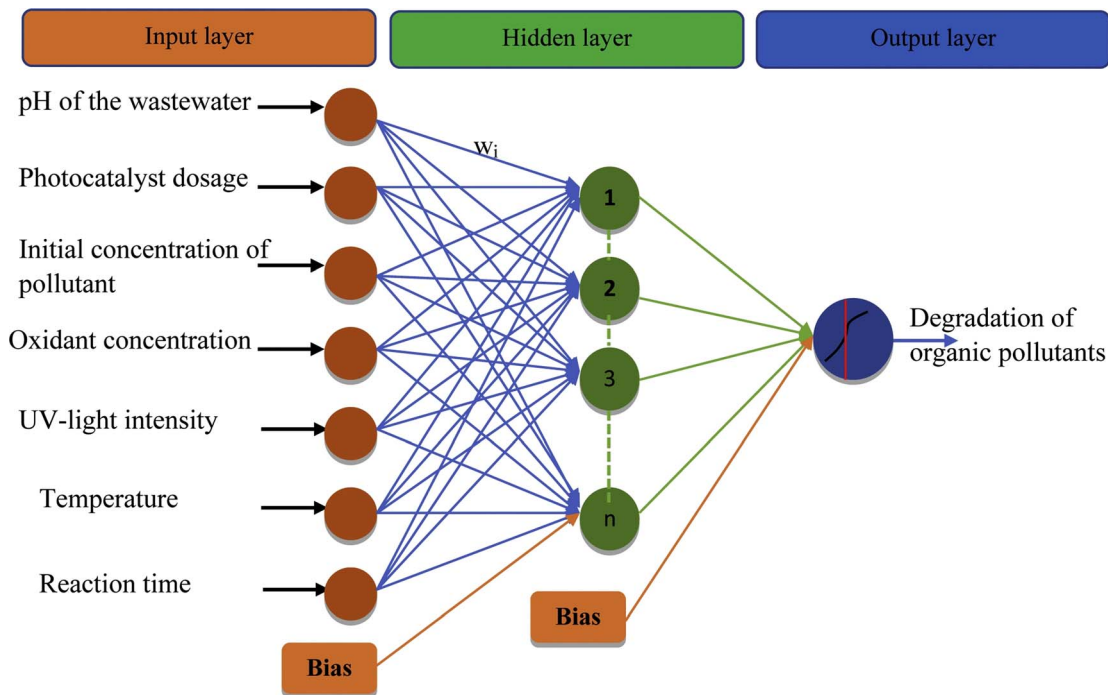
Fig. 5 Setup of an artificial neural network (ANN) for simulating the photocatalytic breakdown of organic pollutants. Copyright (2021) with permission from Elsevier.
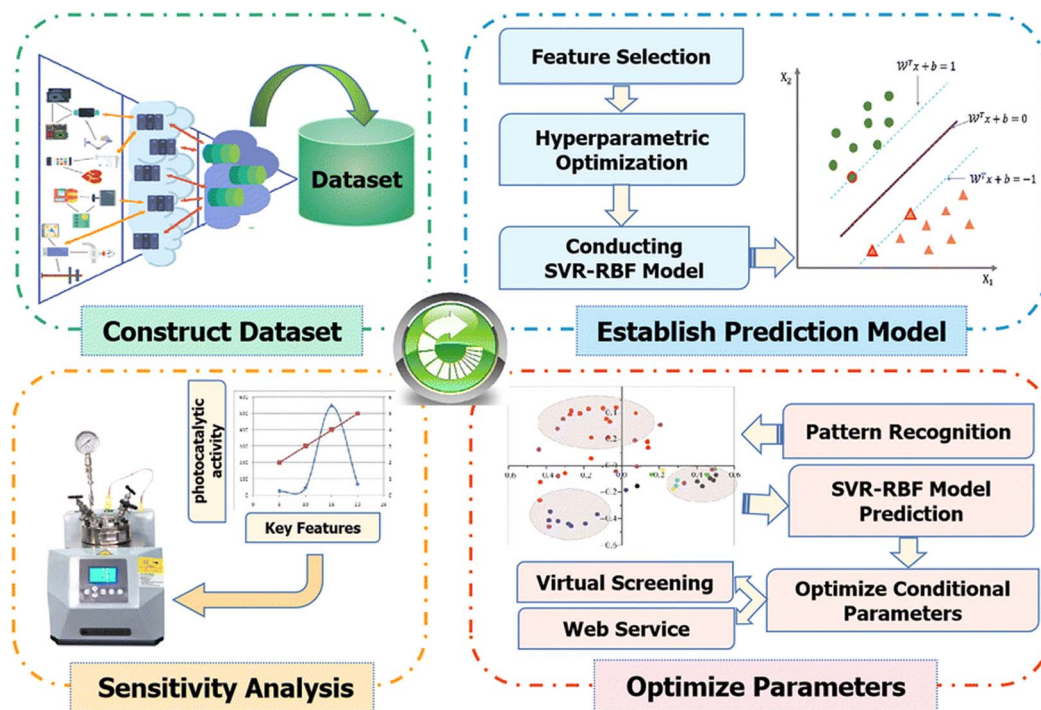

Fig. 6 MLM–BWO/MIL diagram created through a four-step construction process.[31]

testing. The top model successfully captured the dynamics of hydrogen production, as seen by these great results.[150]

Methylene blue was degraded by 94.4% using a BiVO$_4$/BiPO$_4$/rGO nanocatalyst in less than 100 min, according to Yang *et al.* 2023.[151] pH, BiVO$_4$ concentration, GO concentration, ethanol concentration, and volume were the five input variables used in the ANN model (Fig. 7). For neutron numbers greater than 11, the MSE value increased. The MSE values are consistent between 11 and 15. Underfitting occurs when there are too few neutrons, whereas overfitting occurs when there are too many
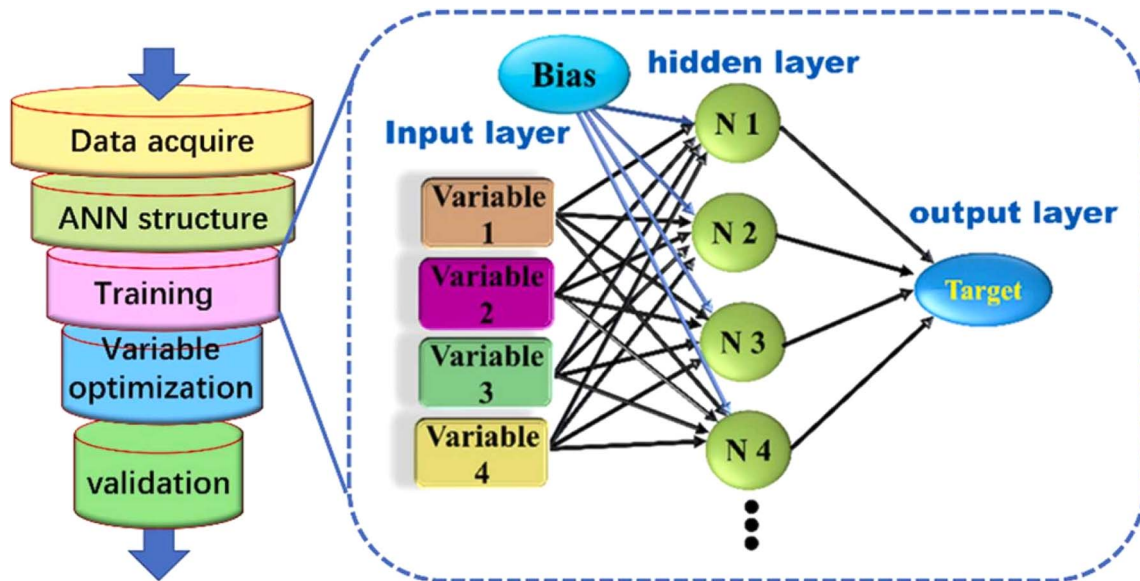
Fig. 7 Schematic representation of the research pathway and structure of an artificial neural network (ANN). Copyright (2023) with permission from Elsevier.

neutrons. This implies that the neutron number of the layer beneath was 11. A scatter plot between the experimental date and the ANN-predicted date, with a slope of the fitted line close to 1, is said to be evidence of the efficacy of the model in the study. They reported overall values of 0.9944, 0.9953, and 1.2040 for the training, test, and validation datasets, respectively. The values of R^2 were 0.9953, 0.9893, and 0.9610. The network performance was improved by experimenting with a wide range of neuron counts in the hidden layer from 2 to 15 (MSE).

There were three sets of data generated at random: three distinct data sets: training, testing, and validation (70 : 15 : 15). It was found that the sparse search algorithm was the best option available. When compared to the experimental value, the ANN model's prediction was only 0.9% off.[151]

5. Challenges and solutions in predicting catalytic degradation performance

One of the challenges in predicting the catalytic degradation performance using machine learning algorithms is understanding the correlation relationships and relative importance of variables. This challenge arises from the complexity and interdependency of the multiple factors that influence the degradation process. To overcome this challenge, researchers can utilize supervised machine-learning approaches to train their models using a dataset that includes a wide range of variables. By incorporating a diverse set of variables, the machine learning algorithm can learn the patterns and relationships between different variables and their impact on catalytic degradation performance.

The availability and quality of the data are other obstacles. Data that are either inaccurate or inadequate can result in

biased predictions and hinder the reliability of machine-learning models. To address this challenge, researchers should ensure that their datasets are comprehensive and representative of the real-world conditions. Data validation procedures should also be used to identify and correct any inconsistencies or inaccuracies in data. Furthermore, the fundamental association between catalyst fingerprint properties and pollutant degradation is difficult to quantify.[40] Researchers can utilize advanced analytical techniques, such as simulations and tests, to quantify the fingerprint properties of catalysts and their impact on pollutant degradation, thus addressing this problem. When combined with experimental data, data-driven machine learning models can help researchers identify natural links between the catalyst properties and the rate at which contaminants in water resources break down. Machine learning's role in predicting catalytic degradation performance machine learning techniques have shown to be important tools for predicting organic pollutant catalytic degradation performance.

Sophisticated algorithms can shift through huge amounts of information in search of patterns and correlations that humans would miss. Scientists can anticipate catalyst features like band gaps and performance measures like adsorption energy and degradation efficiency employing machine learning algorithms.[152]

This enabled them to screen and select potent catalysts for the degradation of organic pollutants more effectively and efficiently. Machine learning models have been found to be particularly useful in the field of catalysis as they can accelerate the process of discovering new catalysts.[152]

These models may be used to analyze catalyst fingerprint parameters to understand their impact on decontamination performance, yield, and reactive oxygen species types. However, it is difficult to identify a quantitative relationship between these fingerprint traits and pollutant degradation. Researchers

are increasingly turning to machine learning approaches to overcome this difficulty. Machine learning models can learn from simulation and experimental data to find intricate correlations between catalyst fingerprint properties and pollutant degradation quantities. Researchers can construct algorithms that accurately predict the catalytic degradation performance of organic pollutants by training these models with large datasets that include information about catalyst characteristics and pollutant degradation performance.

6. Innovations and future developments in machine learning algorithms

One of the main challenges in utilizing machine learning algorithms to estimate the catalytic degradation performance of organic pollutants is recognizing the association between the relative relevance of the components to address this challenge. Ongoing research is focused on exploring advanced feature selection techniques and optimization algorithms to identify the most relevant variables that contribute to the degradation process.[114,115] Another challenge is the complex and nonlinear nature of the catalytic degradation process, which requires the development of more advanced and sophisticated machine learning models.

One example is the use of Explainable Artificial Intelligence (XAI) techniques, together with SHAP (SHapley Additive exPlanations) values, which characterize the contributions of every function to the model's predictions. These methods beautify interpretability, allowing researchers to understand how precise features impact the photocatalytic degradation efficiency.

Researchers are exploring the use of deep learning algorithms, such as neural networks, to capture intricate relationships and patterns in the data. Additionally, advancements in materials science and the integration of machine learning algorithms have opened new possibilities for catalyst discovery. By leveraging data-driven machine learning models, researchers can predict catalyst properties and performance, such as adsorption energy and water-contaminant degradation efficiency.[114,155] These factors play a crucial role in determining the efficiency and selectivity of catalytic reactions. The stoichiometric ratio affects the availability of reactants and can change the rate of the entire reaction, whereas the statistical properties of the elements determine how catalysts react with specific molecules and how well they bind to them. Additionally, the electronic structure of catalysts influences their ability to transfer electrons and participate in redox reactions, thereby affecting the generation and stability of reactive oxygen species.[10,121]

7. Conclusions: future of machine learning in catalytic degradation

The potential of machine learning for predicting the performance of catalytic degradation of organic pollutants is vast and promising. These algorithms have already demonstrated their efficacy in forecasting catalytic rate constants and in enhancing catalyst design. Moreover, machine-learning models hold the key to significantly expediting the development of novel high-performance catalyst materials. Employing machine learning algorithms to analyze vast datasets and unveil intricate patterns and correlations, researchers can gain deeper insight into the mechanisms underlying pollutant degradation. This, in turn, enables the swift screening and selection of effective catalysts, hastening the catalyst development process. By leveraging fingerprint traits, researchers can swiftly identify promising catalyst candidates, thereby reducing their reliance on time-consuming and costly experimentation. In summary, the integration of machine learning algorithms into the realm of catalytic degradation has revolutionized researchers' approaches to pollution remediation. This paradigm shift not only enhances our understanding of the degradation process but also paves the way for faster and more comprehensive pollutant remediation strategies. The field of environmental catalysis stands to benefit significantly from the application of machine learning models, which enable the prediction of organic contaminant degradation efficiency. In conclusion, the utilization of machine learning methods in catalytic degradation research has immense potential for forecasting the degradation of organic pollutants.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 D. Devi Priya, M. d. M. Rahman Khan, S. Mohana Roopan, D. Sreedhar, M. Nandhakumar and M. Ganesapillai, *Int. J. Environ. Anal. Chem.*, 2022, 1–19.

2 D. J. Cram, *Angew. Chem., Int. Ed. Engl.*, 1986, **25**, 1039–1057.

3 R. M. Abu Shmeis, *Compr. Anal. Chem.*, 2022, 105–134.

4 C. Lim and H. Oh, *Annals of Occupational and Environmental Medicine*, 2023, **35**, 1–17.

5 H. Jeong, E. Byeon, D.-H. Kim, P. Maszczyk and J.-S. Lee, *Mar. Pollut. Bull.*, 2023, **191**, 114959.

6 S. Velusamy, A. Roy, S. Sundaram and T. Kumar Mallick, *Chem. Rec.*, 2021, **21**, 1570–1610.

7 K. Samal, S. Mahapatra and M. Hibzur Ali, *Energy Nexus*, 2022, **6**, 100076.

8 M. Berradi, R. Hsissou, M. Khudhair, M. Assouag, O. Cherkaoui, A. El Bachiri and A. El Harfi, *Heliyon*, 2019, **5**, e02711.

9   T. Adane, A. T. Adugna and E. Alemayehu, *J. Chem.*, 2021, **2021**, 1–14.

10 W. A. H. Altowayti, S. Shahir, N. Othman, T. A. E. Eisa, W. M. S. Yafooz, A. Al-Dhaqm, C. Y. Soon, I. B. Yahya, N. A. N. binti Che Rahim, M. Abaker and A. Ali, *Processes*, 2022, **10**, 1832.

11 K. K. Kesari, R. Soni, Q. M. S. Jamal, P. Tripathi, J. A. Lal, N. K. Jha, M. H. Siddiqui, P. Kumar, V. Tripathi and J. Ruokolainen, *Water, Air, Soil Pollut.*, 2021, **232**, 208.

12 X. Zhu, X. Wu, Y. Li, W. Shao, J. Fu, Q. Lin, J. Tan, S. Gao, Y. Zhang and W. Ye, *ACS Appl. Nano Mater.*, 2023, **6**, 5602–5612.

13 S. K. Bhagat, K. E. Pilario, O. E. Babalola, T. Tiyasha, M. Yaqub, C. E. Onu, K. Pyrgaki, M. W. Falah, A. H. Jawad, D. A. Yaseen, N. Barka and Z. M. Yaseen, *J. Clean. Prod.*, 2023, **385**, 135522.

14 G.-W. Cha, H. J. Moon and Y.-C. Kim, *J. Clean. Prod.*, 2022, **375**, 134096.

15 B. Fang, J. Yu, Z. Chen, A. I. Osman, M. Farghali, I. Ihara, E. H. Hamza, D. W. Rooney and P.-S. Yap, *Environ. Chem. Lett.*, 2023, **21**, 1959–1989.

16 G.-W. Cha, S.-H. Choi, W.-H. Hong and C.-W. Park, *Int. J. Environ. Res. Public Health*, 2023, **20**, 3159.

17 Y. Xu, K. Yu, P. Wang, H. Chen, X. Zhao and J. Zhu, *IEEE Access*, 2019, **7**, 138461–138472.

18 M. Abdallah, M. Abu Talib, S. Feroz, Q. Nasir, H. Abdalla and B. Mahfood, *Waste Manag.*, 2020, **109**, 231–246.

19 H. R. Maier, S. Galelli, S. Razavi, A. Castelletti, A. Rizzoli, I. N. Athanasiadis, M. Sànchez-Marrè, M. Acutis, W. Wu and G. B. Humphrey, *Environ. Model. Software*, 2023, **167**, 105776.

20 A. T. Taiba Mustafa, A. Ghulam and A. Akram, *J. Jilin Univ.*, 2023, **42**, 386–406.

21 R. Kessels and G. Erreygers, *Health Econ.*, 2019, **28**, 884–905.

22 V. G. Costa and C. E. Pedreira, *Artif. Intell. Rev.*, 2023, **56**, 4765–4800.

23 P. S. Sajja, Examples and applications on genetic algorithms, *Illustrated Computational Intelligence*, 2021, pp. 155–189.

24 X. Liu, D. Lu, A. Zhang, Q. Liu and G. Jiang, *Environ. Sci. Technol.*, 2022, **56**, 2124–2133.

25 S. Geetha, J. Saha, I. Dasgupta, R. Bera, I. A. Lawal and S. Kadry, *Designs*, 2022, **6**, 27.

26 S. Vyas, K. Dhakar, S. Varjani, R. R. Singhania, P. C. Bhargava, R. Sindhu, P. Binod, J. W. C. Wong and X.-T. Bui, *Sci. Total Environ.*, 2023, **891**, 164344.

27 J. A. Silva, *Sustainability*, 2023, **15**, 10940.

28 K. Rasouli, A. Alamdari and S. Sabbaghi, *Sep. Purif. Technol.*, 2023, **307**, 122799.

29 J. Malinauskaite, H. Jouhara, D. Czajczyńska, P. Stanchev, E. Katsou, P. Rostkowski, R. J. Thorne, J. Colón, S. Ponsá, F. Al-Mansour, L. Anguilano, R. Krzyżyńska, I. C. López, A. Vlasopoulos and N. Spencer, *Energy*, 2017, **141**, 2013–2044.

30 F.-C. Mihai, S. Gündoğdu, L. A. Markley, A. Olivelli, F. R. Khan, C. Gwinnett, J. Gutberlet, N. Reyna-Bensusan, P. Llanquileo-Melgarejo, C. Meidiana, S. Elagroudy, V. Ishchenko, S. Penney, Z. Lenkiewicz and M. Molinos-Senante, *Sustainability*, 2021, **14**, 20.

31 X. Zhai and M. Chen, *Nanoscale Adv.*, 2023, **5**, 4065–4073.

32 A. S. Ahuja, *PeerJ*, 2019, **7**, e7702.

33 V. K. Pandey, S. Srivastava, K. K. Dash, R. Singh, S. A. Mukarram, B. Kovács and E. Harsányi, *Processes*, 2023, **11**, 1720.

34 D. Paul, G. Sanap, S. Shenoy, D. Kalyane, K. Kalia and R. K. Tekade, *Drug Discovery Today*, 2021, **26**, 80–93.

35 K. Pouthika and G. Madhumitha, *Inorg. Chim. Acta*, 2023, **551**, 121457.

36 G. Wang, S. Zong, H. Ma, B. Wan and Q. Tian, *Catalysts*, 2023, **13**, 184.

37 A. Saravanan, P. Senthil Kumar, S. Jeevanantham, S. Karishma, B. Tajsabreen, P. R. Yaashikaa and B. Reshma, *Chemosphere*, 2021, **280**, 130595.

38 M. Nageeb, in *Organic Pollutants – Monitoring, Risk and Treatment*, InTech, 2013.

39 H.-Y. Shyu, C. J. Castro, R. A. Bair, Q. Lu and D. H. Yeh, *ACS Environ. Au*, 2023, **3**, 308–318.

40 A. El Jery, M. Aldrdery, U. R. Shirode, J. C. O. Gavilán, A. Elkhaleefa, M. Sillanpää, S. S. Sammen and H. H. Tizkam, *Catalysts*, 2023, **13**, 1085.

41 F. Granata, S. Papirio, G. Esposito, R. Gargano and G. De Marinis, *Water*, 2017, **9**, 105.

42 W. Cai, F. Long, Y. Wang, H. Liu and K. Guo, *Microb. Biotechnol.*, 2021, **14**, 59–62.

43 M. Bahramian, R. K. Dereli, W. Zhao, M. Giberti and E. Casey, *Expert Syst. Appl.*, 2023, **217**, 119453.

44 F. Gao, Y. Shen, J. Brett Sallach, H. Li, W. Zhang, Y. Li and C. Liu, *J. Hazard. Mater.*, 2022, **424**, 127437.

45 A. I. Osman, Y. Zhang, Z. Y. Lai, A. K. Rashwan, M. Farghali, A. A. Ahmed, Y. Liu, B. Fang, Z. Chen, A. Al-Fatesh, D. W. Rooney, C. L. Yiin and P.-S. Yap, *Environ. Chem. Lett.*, 2023, **21**, 3159–3244.

46 H. Masood, C. Y. Toe, W. Y. Teoh, V. Sethu and R. Amal, *ACS Catal.*, 2019, **9**, 11774–11787.

47 J. C. Y. Ngu, M. K. Chan, W. S. Yeo and J. Nandong, *MATEC Web Conf.*, 2023, **377**, 01009.

48 J. Ohyama, S. Nishimura and K. Takahashi, *ChemCatChem*, 2019, **11**, 4307–4313.

49 S. Revollar, P. Vega, R. Vilanova and M. Francisco, *Appl. Sci.*, 2017, **7**, 813.

50 P. Rajasulochana and V. Preethy, *Resour.-Effic. Technol.*, 2016, **2**, 175–184.

51 S. Mohana Roopan and M. A. Khan, *Catal. Rev.*, 2023, **65**, 620–693.

52 W. Xia, Y. Jiang, X. Chen and R. Zhao, *Waste Manag. Res.*, 2022, **40**, 609–624.

53 R. Khan, S. Kumar, A. K. Srivastava, N. Dhingra, M. Gupta, N. Bhati and P. Kumari, *Comput. Intell. Neurosci.*, 2021, **2021**, 1–11.

54 M. T. Munir, B. Li and M. Naqvi, *Fuel*, 2023, **348**, 128548.

55 S. D. Apte, S. Sandbhor, R. Kulkarni and H. Khanum, *Front. Mech. Eng.*, 2023, **9**, 1–11.

56 S. Chaturvedi, B. P. Yadav and N. A. Siddiqui, *Nat., Environ. Pollut. Technol.*, 2021, **20**, 1515–1525.

57 G. K. V. Nachiyar, T. V. Surendra, V. Kalaiselvi, R. Rajagopal, P. Kuppusamy, N. Basavegowda and S. M. Roopan, *Optik*, 2022, **267**, 169633.

58 H. Guo, S. Wu, Y. Tian, J. Zhang and H. Liu, *Bioresour. Technol.*, 2021, **319**, 124114.

59 P. Sharma and U. Vaid, *IOP Conf. Ser. Earth Environ. Sci.*, 2021, **889**, 012047.

60 S. Rubab, M. M. Khan, F. Uddin, Y. Abbas Bangash and S. A. A. Taqvi, *ChemBioEng Rev.*, 2022, **9**, 212–226.

61 E. Mati Asefa, K. Bayu Barasa and D. Adare Mengistu, in *Geographic Information Systems and Applications in Coastal Studies*, IntechOpen, 2022.

62 V. Chang, J. Bailey, Q. A. Xu and Z. Sun, *Neural Comput. Appl.*, 2023, **35**, 16157–16173.

63 L. Benos, A. C. Tagarakis, G. Dolias, R. Berruto, D. Kateris and D. Bochtis, *Sensors*, 2021, **21**, 3758.

64 A. H. Saad, H. Nahazanan, B. Yusuf, S. F. Toha, A. Alnuaim, A. El-Mouchi, M. Elseknidy and A. A. Mohammed, *Sustainability*, 2023, **15**, 9738.

65 D. Liu and N. Sun, *Frontiers in Sustainability*, 2023, **4**, 1–8.

66 J. Behera, A. K. Pasayat, H. Behera and P. Kumar, *Eng. Appl. Artif. Intell.*, 2023, **120**, 105843.

67 S. Abu Naser, A. M. H. Taha and D. Ariffin, *J. Theor. Appl. Inf. Technol.*, 2023, **101**, 21–36.

68 P. Carracedo-Reboredo, J. Liñares-Blanco, N. Rodríguez-Fernández, F. Cedrón, F. J. Novoa, A. Carballal, V. Maojo, A. Pazos and C. Fernandez-Lozano, *Comput. Struct. Biotechnol. J.*, 2021, **19**, 4538–4558.

69 D. A. Hashimoto, E. Witkowski, L. Gao, O. Meireles and G. Rosman, *Anesthesiology*, 2020, **132**, 379–394.

70 M. M. Taye, *Computers*, 2023, **12**, 91.

71 P. Linardatos, V. Papastefanopoulos and S. Kotsiantis, *Entropy*, 2020, **23**, 18.

72 P. Kumar, *International Journal of Data Mining & Knowledge Management Process*, 2012, **2**, 25–42.

73 S. F. Ahmed, M. S. Bin Alam, M. Hassan, M. R. Rozbu, T. Ishtiak, N. Rafa, M. Mofijur, A. B. M. Shawkat Ali and A. H. Gandomi, *Artif. Intell. Rev.*, 2023, **56**, 13521–13617.

74 J. Shen and M. O. Shafiq, *Journal of Big Data*, 2020, **7**, 66.

75 B. A. Marinho, L. Suhadolnik, B. Likozar, M. Huš, Ž. Marinko and M. Čeh, *J. Clean. Prod.*, 2022, **343**, 131061.

76 M. Lowe, R. Qin and X. Mao, *Water*, 2022, **14**, 1384.

77 B. Abebe, H. C. A. Murthy and E. Amare, *J. Encapsulation Adsorpt. Sci.*, 2018, **08**, 225–255.

78 S. Mitra and G. S. Murthy, *Syst. Microbiol. Biomanuf.*, 2022, **2**, 91–112.

79 Z. Ye, J. Yang, N. Zhong, X. Tu, J. Jia and J. Wang, *Sci. Total Environ.*, 2020, **699**, 134279.

80 P. Schlexer Lamoureux, K. T. Winther, J. A. Garrido Torres, V. Streibel, M. Zhao, M. Bajdich, F. Abild-Pedersen and T. Bligaard, *ChemCatChem*, 2019, **11**, 3581–3601.

81 M. Alshurideh, B. Al Kurdi, S. A. Salloum, I. Arpaci and M. Al-Emran, *Interact. Learn. Environ.*, 2023, **31**, 1214–1228.

82 B. V. Ayodele, M. A. Alsaffar, S. I. Mustapa and D. N. Vo, *J. Chem. Technol. Biotechnol.*, 2020, **95**, 2739–2749.

83 N. Kitchamsetti and A. L. F. de Barros, *ChemCatChem*, 2023, **15**, e202300690.

84 J. Bao, S. Guo, D. Fan, J. Cheng, Y. Zhang and X. Pang, *Ultrason. Sonochem.*, 2023, **99**, 106569.

85 M. Gheytanzadeh, A. Baghban, S. Habibzadeh, K. Jabbour, A. Esmaeili, A. Mohaddespour and O. Abida, *Sci. Rep.*, 2022, **12**, 6615.

86 M. Umar and H. Abdul, in *Organic Pollutants – Monitoring, Risk and Treatment*, InTech, 2013.

87 M. Al-Emran, M. N. Al-Nuaimi, I. Arpaci, M. A. Al-Sharafi and B. Anthony Jr, *Educ. Inf. Technol.*, 2023, **28**, 2727–2746.

88 M. S. Zaghloul and G. Achari, *J. Environ. Chem. Eng.*, 2022, **10**, 107430.

89 S. Radović, S. Pap and M. Turk Sekulić, in *Proceedings – the Eleventh International Symposium GRID 2022*, University of Novi Sad, Faculty of Technical Sciences, Department of Graphic Engineering and Design, 2022, pp. 799–807.

90 M. Li, K. Hu and J. Wang, *J. Eng. Appl. Sci.*, 2021, **68**, 35.

91 M. A. Alsaffar, M. A. R. A. Ghany, A. K. Mageed, A. A. AbdulRazak, J. M. Ali, K. A. Sukkar and B. V. Ayodele, *Appl. Sci.*, 2023, **13**, 8966.

92 R. Mohammadzadeh Kakhki, S. Zirjanizadeh and M. Mohammadpoor, *J. Mater. Sci.*, 2023, **58**, 10555–10575.

93 S. Yadav, K. Shakya, A. Gupta, D. Singh, A. R. Chandran, A. Varayil Aanappalli, K. Goyal, N. Rani and K. Saini, *Environ. Sci. Pollut. Res.*, 2022, **30**, 71912–71932.

94 L. Zhao, T. Dai, Z. Qiao, P. Sun, J. Hao and Y. Yang, *Process Saf. Environ. Prot.*, 2020, **133**, 169–182.

95 Z. Frontistis, G. Lykogiannis and A. Sarmpanis, *Environments*, 2023, **10**, 127.

96 B. Sundui, O. A. Ramirez Calderon, O. M. Abdeldayem, J. Lázaro-Gil, E. R. Rene and U. Sambuu, *Clean Technol. Environ. Policy*, 2021, **23**, 127–143.

97 G. Ren, H. Han, Y. Wang, S. Liu, J. Zhao, X. Meng and Z. Li, *Nanomaterials*, 2021, **11**, 1804.

98 S. Martini and K. A. Roni, *J. Phys.: Conf. Ser.*, 2021, **1858**, 012013.

99 A. Pattnaik, J. N. Sahu, A. K. Poonia and P. Ghosh, *Chem. Eng. Res. Des.*, 2023, **190**, 667–686.

100 M. Nasr, K. Mohamed, M. Attia and M. G. Ibrahim, in *Soft Computing Techniques in Solid Waste and Wastewater Management*, Elsevier, 2021, pp. 171–185.

101 N. S. Singh, C. Murugamani, P. R. Kshirsagar, V. Tirth, S. Islam, S. Qaiyum, B. Suneela, M. Al Duhayyim and Y. A. Waji, *Sci. Program.*, 2022, **2022**, 1–11.

102 S. N. Zulkifli, H. A. Rahim and W.-J. Lau, *Sens. Actuators, B*, 2018, **255**, 2657–2689.

103 J. Y. Uwamungu, P. Kumar, A. Alkhayyat, T. Younas, R. Y. Capangpangan, A. C. Alguno and I. Ofori, *J. Nanomater.*, 2022, **2022**, 1–11.

104 N. M. Hosny, I. Gomaa and M. G. Elmhgary, *Applied Surface Science Advances*, 2023, **15**, 100395.

105 P. Ramesh and A. Rajendran, *Chemical Physics Impact*, 2023, **6**, 100208.

106 S. S. Gill, M. Xu, C. Ottaviani, P. Patros, R. Bahsoon, A. Shaghaghi, M. Golec, V. Stankovski, H. Wu, A. Abraham, M. Singh, H. Mehta, S. K. Ghosh, T. Baker, A. K. Parlikad, H. Lutfiyya, S. S. Kanhere, R. Sakellariou,

S. Dustdar, O. Rana, I. Brandic and S. Uhlig, *Internet of Things*, 2022, **19**, 100514.

107 S. Khan, M. Naushad, M. Govarthanan, J. Iqbal and S. M. Alfadul, *Environ. Res.*, 2022, **207**, 112609.

108 X. Fu, Y. Han, H. Xu, Z. Su and L. Liu, *J. Hazard. Mater.*, 2022, **422**, 126890.

109 A. R. Dash, A. J. Lakhani, D. Devi Priya, T. V. Surendra, M. M. R. Khan, E. J. J. Samuel and S. M. Roopan, *J. Cluster Sci.*, 2023, **34**, 121–133.

110 H. Ruan, J. Chen, W. Ai and B. Wu, *Energy and AI*, 2022, **9**, 100158.

111 A. Jamal, I. Reza and M. Shafiullah, *IATSS Res.*, 2022, **46**, 499–514.

112 A. M. Ghaedi and A. Vafaei, *Adv. Colloid Interface Sci.*, 2017, **245**, 20–39.

113 M. Fan, J. Hu, R. Cao, W. Ruan and X. Wei, *Chemosphere*, 2018, **200**, 330–343.

114 S. Liu, C. K. Y. Lo and C. Kan, *Color. Technol.*, 2022, **138**, 117–136.

115 S. K. Bhagat, K. E. Pilario, O. E. Babalola, T. Tiyasha, M. Yaqub, C. E. Onu, K. Pyrgaki, M. W. Falah, A. H. Jawad, D. A. Yaseen and N. Barka, *J. Cleaner Prod.*, 2023, **385**, 1–64.

116 N. Taoufik, W. Boumya, M. Achak, H. Chennouk, R. Dewil and N. Barka, *Sci. Total Environ.*, 2022, **807**, 150554.

117 A. H. Navidpour, A. Hosseinzadeh, Z. Huang, D. Li and J. L. Zhou, *Catal. Rev.*, 2022, 1–26.

118 V. Godvin Sharmila, V. Kumar Tyagi, S. Varjani and J. Rajesh Banu, *Bioresour. Technol.*, 2023, **387**, 129587.

119 P. Ifaei, M. Nazari-Heris, A. S. Tayerani Charmchi, S. Asadi and C. Yoo, *Energy*, 2023, **266**, 126432.

120 P. Manickam, S. A. Mariappan, S. M. Murugesan, S. Hansda, A. Kaushik, R. Shinde and S. P. Thipperudraswamy, *Biosensors*, 2022, **12**, 562.

121 T.-N. Do, D.-M. T. Nguyen, J. Ghimire, K.-C. Vu, L.-P. Do Dang, S.-L. Pham and V.-M. Pham, *Environ. Sci. Pollut. Res.*, 2023, **30**, 82230–82247.

122 H. Mai, T. C. Le, D. Chen, D. A. Winkler and R. A. Caruso, *Chem. Rev.*, 2022, **122**, 13478–13515.

123 V. E. Sathishkumar, A. G. Ramu and J. Cho, *Alexandria Eng. J.*, 2023, **72**, 673–693.

124 S. S. Jeon, P. W. Kang, M. Klingenhof, H. Lee, F. Dionigi and P. Strasser, *ACS Catal.*, 2023, **13**, 1186–1196.

125 Y. Su, S. Wang, L. Ji, C. Zhang, H. Cai, H. Zhang and W. Sun, *Nanoscale*, 2023, **15**, 154–161.

126 B. Dong, N. Xue, G. Mu, M. Wang, Z. Xiao, L. Dai, Z. Wang, D. Huang, H. Qian and W. Chen, *Ultrason. Sonochem.*, 2021, **73**, 105485.

127 J. S. Dondapati and A. Chen, *Phys. Chem. Chem. Phys.*, 2020, **22**, 8878–8888.

128 T. Chellapandi and G. Madhumitha, *Environ. Qual. Manag.*, 2023, DOI: 10.1002/tqem.22021.

129 H. H. Shanaah, E. F. H. Alzaimoor, S. Rashdan, A. A. Abdalhafith and A. H. Kamel, *Sustainability*, 2023, **15**, 7336.

130 O. A. Moses, W. Chen, M. L. Adam, Z. Wang, K. Liu, J. Shao, Z. Li, W. Li, C. Wang, H. Zhao, C. H. Pang, Z. Yin and X. Yu, *Mater. Rep.: Energy*, 2021, **1**, 100049.

131 M. Joy, G. Chandrasekharan, M. A. Khan, M. Arunachalapandi, T. Chellapandi, D. Harish, D. Chitra and S. M. Roopan, *Environ. Qual. Manag.*, 2022, **32**, 159–169.

132 D. Yu, F. Wu, J. He, L. Bai, Y. Zheng, Z. Wang and J. Zhang, *Appl. Catal., B*, 2023, **320**, 121880.

133 S. Sarmin, M. Tarek, C. K. Cheng, S. M. Roopan and M. M. R. Khan, *J. Hazard. Mater.*, 2021, **415**, 125587.

134 S. Shanavas, S. Mohana Roopan, A. Priyadharsan, D. Devipriya, S. Jayapandi, R. Acevedo and P. M. Anbarasan, *Appl. Catal., B*, 2019, **255**, 117758.

135 K. Anand, K. Kaviyarasu, S. Muniyasamy, S. M. Roopan, R. M. Gengan and A. A. Chuturgoon, *J. Cluster Sci.*, 2017, **28**, 2279–2291.

136 J. Abdi, M. Hadipoor, F. Hadavimoghaddam and A. Hemmati-Sarapardeh, *Chemosphere*, 2022, **287**, 132135.

137 W. Liang, T. Liu, Y. Wang, J. J. Jiao, J. Gan and D. He, *Sci. Total Environ.*, 2023, **905**, 167138.

138 S. D. Shelare, P. N. Belkhode, K. C. Nikam, L. D. Jathar, K. Shahapurkar, M. E. M. Soudagar, I. Veza, T. M. Y. Khan, M. A. Kalam, A.-S. Nizami and M. Rehan, *Energy*, 2023, **282**, 128874.

139 N. J. Aquilina, J. M. Delgado-Saborit, S. Bugelli, J. P. Ginies and R. M. Harrison, *Environ. Sci. Technol.*, 2018, **52**, 11215–11222.

140 R. Rakholia, Q. Le, B. Quoc Ho, K. Vu and R. Simon Carbajo, *Environ. Int.*, 2023, **173**, 107848.

141 J. Lin, F. Gulbagca, A. Aygun, R. N. Elhouda Tiri, C. Xia, Q. Van Le, T. Gur, F. Sen and Y. Vasseghian, *Food Chem. Toxicol.*, 2022, **163**, 112972.

142 C. Bellinger, M. S. Mohomed Jabbar, O. Zaïane and A. Osornio-Vargas, *BMC Publ. Health*, 2017, **17**, 907.

143 B. V. Ayodele, M. A. Alsaffar, S. I. Mustapa, C. K. Cheng and T. Witoon, *Process Saf. Environ. Prot.*, 2021, **145**, 120–132.

144 A. Gordanshekan, S. Arabian, A. R. Solaimany Nazar, M. Farhadian and S. Tangestaninejad, *Chem. Eng. J.*, 2023, **451**, 139067.

145 O. Hosseini, V. Zare-Shahabadi, M. Ghaedi and M. H. A. Azghandi, *J. Environ. Chem. Eng.*, 2022, **10**, 108345.

146 Q. Shang, X. Liu, M. Zhang, P. Zhang, Y. Ling, G. Cui, W. Liu, X. Shi, J. Yue and B. Tang, *Chem. Eng. J.*, 2022, **443**, 136354.

147 Z. H. Jaffari, A. Abbas, S.-M. Lam, S. Park, K. Chon, E.-S. Kim and K. H. Cho, *J. Hazard. Mater.*, 2023, **442**, 130031.

148 N. M. Mahmoodi, J. Abdi, M. Taghizadeh, A. Taghizadeh, B. Hayati, A. A. Shekarchi and M. Vossoughi, *J. Environ. Manage.*, 2019, **233**, 660–672.

149 J. Rodrigues, T. Hatami, J. M. Rosa, E. B. Tambourgi and L. H. I. Mei, *Chem. Eng. Res. Des.*, 2020, **153**, 294–305.

150 J. Li, Z. Chen, J. Wu, J. Lin, P. He, R. Zhu, C. Peng, H. Zhang, W. Li, X. Fang and H. Shen, *Mater. Today Commun.*, 2023, **35**, 106299.

© 2024 The Author(s). Published by the Royal Society of Chemistry

151 F. Yang, X. Yu, K. Wang, Z. Liu, Z. Gao, T. Zhang, J. Niu, J. Zhao and B. Yao, *J. Alloys Compd.*, 2023, **960**, 170716.

152 Z. Jiang, J. Hu, A. Samia and X. Yu, *Catalysts*, 2022, **12**, 746.

153 P. Venkatasaichandrakanth and M. Iyapparaja, *IEEE Access*, 2023, **11**, 54045–54057.

154 P. Venkatasaichandrakanth and M. Iyapparaja, in *Artificial Intelligence and Machine Learning for Smart Communities*, CRC Press, 2024, pp. 1–25.

155 S. Gautam, H. Agrawal, M. Thakur, A. Akbari, H. Sharda, R. Kaur and M. Amini, *J. Environ. Chem. Eng.*, 2020, 103726.

156 R. Ameta and S. C. Ameta, *Photocatalysis: Principles and Applications*, Crc Press, 2016.

157 R. M. Zaki, R. Jusoh, I. Chanakaewsomboon and H. D. Setiabudi, *Mater. Today: Proc.*, 2023, 1–9.

158 K. M. Mohamed, J. J. Benitto, J. J. Vijaya and M. Bououdina, *Crystals*, 2023, **13**, 329.

159 A. Rafiq, M. Ikram, S. Ali, F. Niaz, M. Khan, Q. Khan and M. Maqbool, *J. Ind. Eng. Chem.*, 2021, **97**, 111–128.

160 M. A. Fox and M. T. Dulay, *Chem. Rev.*, 1993, **93**, 341–357.

161 K. M. Reza, A. S. W. Kurny and F. Gulshan, *Appl. Water Sci.*, 2017, **7**, 1569–1578.

162 H. M. Coleman, V. Vimonses, G. Leslie and R. Amal, *J. Hazard. Mater.*, 2007, **146**, 496–501.

163 U. I. Gaya and A. H. Abdullah, *J. Photochem. Photobiol., C*, 2008, **9**, 1–12.

164 L. Karimi, M. E. Yazdanshenas, R. Khajavi, A. Rashidi and M. Mirjalili, *Appl. Surf. Sci.*, 2015, **332**, 665–673.