

Article

Interpretable Machine Learning Approach for Photocatalytic Degradation in Mn-Doped Semiconductors Using Multilayer Perceptron and SHAP Analysis

Orhan Baytar ¹, Metin Zontul ^{2,*} , Ceren Orak ³, Seda Karateke ⁴ , Hakan Aydın ⁵  and Sabit Horoz ^{6,7,*} ¹ Department of Chemical Engineering, Siirt University, 56100 Siirt, Türkiye; orhan.baytar@siirt.edu.tr² Department of Computer Engineering, Sivas University of Science and Technology, 58140 Sivas, Türkiye³ Department of Chemical Engineering, Sivas University of Science and Technology, 58140 Sivas, Türkiye; cerenorak@sivas.edu.tr⁴ Department of Software Engineering, Istanbul Atlas University, 34408 İstanbul, Türkiye; seda.karateke@atlas.edu.tr⁵ Department of Computer Engineering, İstanbul Topkapı University, 34087 İstanbul, Türkiye; hakanaydin@topkapi.edu.tr⁶ Department of Engineering Fundamental Sciences, Sivas University of Science and Technology, 58140 Sivas, Türkiye⁷ Nanophotonics Application and Research Center-CÜNAM, Sivas Cumhuriyet University, 58140 Sivas, Türkiye

* Correspondence: metinzontul@sivas.edu.tr (M.Z.); sabit.horoz@sivas.edu.tr (S.H.)

Abstract

This study comprehensively investigates the degradation performance of a Mn-doped Zn_2SnO_4 photocatalyst based on time-dependent UV-Vis absorption spectra. Before machine learning modelling, the effects of experimental parameters such as UV-Vis measurement wavelength, reaction time, and Mn doping ratio were statistically validated using One-Way Analysis of Variance (ANOVA) and Multiple Linear Regression (MLR) methods. To overcome the limitations of linear models in representing complex physical systems, an optimized Multi-Layer Perceptron (MLP) architecture was developed to capture the system's nonlinear dynamics with high accuracy. To validate the model's out-of-sample prediction capability and prevent data leakage potentially arising from spectral data correlation, the "Leave-One-Doping-Level-Out" (LODLO) cross-validation strategy was applied, during which performance metrics of $R^2 = 0.8889$ and $MSE = 0.00238$ were recorded. To make the neural network's decision-making mechanism transparent, a dual-validation explainability framework comprising Shapley Additive Explanations (SHAP) and Permutation Feature Importance analyses was employed. By quantifying the relative contributions of the experimental parameters to the model predictions, this approach revealed that the UV-Vis measurement wavelength was the dominant predictive variable, followed by the Mn doping ratio and reaction time. This study presents a transparent methodology that offers both strong predictive capability and physically grounded data to shed light on the complex interactions in doped semiconductor photocatalysts.

Keywords: Mn-substituted Zn_2SnO_4 ; neural network modelling; LODLO cross-validation; Explainable AI (XAI); photodegradation kinetics; analysis of variance (ANOVA); UV-Vis Spectral Dynamics



Academic Editor: Albert Poater

Received: 10 April 2026

Revised: 13 May 2026

Accepted: 15 May 2026

Published: 8 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Visible-light photocatalysis has emerged as a broadly applicable platform technology that encompasses two fundamentally distinct paradigms. The first involves molecular organophotocatalysis, in which photoredox-active organic dyes or transition metal complexes drive selective bond-forming reactions under mild conditions, as exemplified by recent advances in visible-light-mediated organic synthesis [1]. The second, and the focus of the present work, is heterogeneous semiconductor photocatalysis, in which photo-generated electron–hole pairs at the surface of an inorganic semiconductor drive the oxidative mineralization of environmental pollutants. Although both fields share the common principle of light-driven activation, they differ fundamentally in their catalytic architecture, reaction mechanism, and application target. The present study is situated exclusively within the latter paradigm, investigating the photocatalytic degradation of methylene blue over Mn-doped Zn_2SnO_4 under UV irradiation.

Zinc stannate (Zn_2SnO_4), a ternary metal oxide with an inverse spinel structure, is a material that has attracted attention in environmental remediation efforts due to its high electron mobility and thermodynamic stability [2,3]. However, the material's wide bandgap in the 3.6–3.8 eV range limits its practical applications by restricting its ability to utilize only the ultraviolet portion of the solar spectrum [4]. Additionally, the rapid recombination kinetics of photo-excited electron–hole pairs constitute a fundamental barrier that suppresses quantum efficiency [5]. To overcome these bottlenecks, transition metal doping emerges as a strategic solution. In particular, manganese (Mn) [6–8] can be successfully integrated into the Zn_2SnO_4 matrix due to its half-filled 3d electron configuration, creating intermediate energy levels within the bandgap and thereby extending the lifetime of charge carriers [9–11].

Although it is known that Mn doping at low and medium concentrations increases the production of reactive oxygen species (ROS) [11–14], the relationship between the doping concentration and photocatalytic efficiency is not always linear. Once a critical threshold is exceeded, excess Mn ions can cause distortions in the crystal lattice, turning it into sites where charges recombine [15]. Such complex and nonlinear behaviours necessitate analytical tools that go beyond traditional trial-and-error methods or classical statistical models such as One-Factor-At-a-Time (OFAT) and Response Surface Methodology (RSM) [16,17]. In recent years, the paradigm in materials science has shifted toward data-driven machine learning (ML) architectures [18,19]. Deep learning algorithms such as MLPs and Support Vector Regression (SVR) have the capacity to map complex physical phenomena without the need for predefined parametric equations [20–22]. However, the high predictive power of these models brings with it the “black-box” dilemma [23]; the physical and chemical mechanisms underlying the model's results often remain unclear. In the field of materials chemistry, predictive success alone is insufficient; understanding the driving forces behind these predictions is essential for rational catalyst design [24,25]. For instance, a comprehensive review by Armačković et al. emphasized that integrating machine learning techniques with ZnO-based photocatalyst design can significantly accelerate the optimization of synthesis parameters, enhance charge transport understanding, and improve overall photocatalytic performance [26].

To address this gap, the integration of Explainable Artificial Intelligence (XAI) has become a necessity. The SHAP algorithm, based on cooperative game theory, ensures transparency by mathematically calculating the marginal contribution of each input feature to the model's output [27,28]. A review of the literature reveals that while the applications of machine learning in binary oxides have increased, the predictive modelling of complex ternary systems, such as Mn-doped Zn_2SnO_4 , remains largely unexplored. More

importantly, existing studies typically stop at the prediction stage and do not translate their findings into mechanistic insights via XAI [29,30].

This study was designed to overcome this methodological challenge. By combining an optimized MLP framework with a dual-validation explainability approach incorporating both kernel-based SHAP analysis and Permutation Feature Importance, we aim to predict the degradation dynamics of Mn-doped Zn_2SnO_4 and interpret the learned associations between the UV–Vis measurement wavelength, reaction time, and Mn doping ratio. The Permutation Feature Importance analysis was included as an independent, model-agnostic cross-validation of the SHAP results, ensuring that the feature importance hierarchy is not an artefact of a single explainability method but is reproducible across different mathematical frameworks. In addition, pseudo-first-, pseudo-second-, and pseudo-third-order kinetic analyses were performed on the time-resolved absorbance data to rigorously characterise the reaction mechanism governing MB degradation, determine the apparent rate constants across all doping levels, and identify the kinetic model that best describes the experimental observations. By integrating these analytical layers, we depart from classical data fitting methods to provide a transparent, physically interpretable, and kinetically grounded roadmap for the optimisation of transition-metal-doped photocatalytic systems.

2. Results and Discussion

2.1. Prediction Strategy Using MLP Regression

To overcome the limitations of linear models, we utilized an MLP architecture to model the absorption data of the photocatalytic process [31] which exhibits a complex, time-varying nature. Our primary objective in designing this network architecture (Figure 1) was to accurately map the multidimensional effects of parameters such as UV–Vis measurement wavelength, reaction time, and Mn loading, which carry distinct physical meanings, on the target output. However, during the machine learning process, the inherent scale differences in the raw data can destabilize the gradient descent algorithm, potentially reducing the convergence rate. To prevent such computational bottlenecks and stabilize the model's learning performance, we pre-processed the data using standard scaling [32]. While alternative transformations, such as wavelet or Fourier transforms [33], might be preferred in certain specific scenarios, this normalization procedure proved sufficient for the current dataset. This mathematical transformation calculates the scaled z values using the following formula:

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

Here, x represents the raw observation, μ denotes the mean of the dataset, and σ represents the standard deviation. This preparatory step ensures that experimental variables with different units are represented in a balanced manner within the model.

In addition, the dataset was randomly divided into training (60%) and test (40%) subsets for algorithmic evaluation. The dataset consists of a total of 3660 experimental observations, generated through a full-factorial design encompassing all combinations of the three independent variables. The detailed experimental plan is presented in Table 1.

Each experimental condition was independently repeated three times to ensure measurement reproducibility, and the reported absorbance values correspond to the averaged results of these repeated trials. This experimental replication protocol effectively minimizes random measurement noise and enhances the statistical robustness of the dataset. Furthermore, the dataset provides a well-distributed representation of the parameter space, ensuring that the trained model captures the underlying physicochemical trends without bias toward specific operating conditions. To ensure the consistency and relia-

bility of the dataset, all experimental data in the study were generated in a controlled laboratory environment.

Table 1. Full-factorial experimental design showing the independent variables and their corresponding levels used in this study.

Parameter	Unit	Number of Levels	Levels
UV–Vis Measurement Wavelength	nm	61	200, 210, 220, . . . , 800 (10 nm increments)
Reaction Time	min	10	0, 10, 20, 30, 40, 50, 60, 70, 80, 90
Mn Doping Ratio	%	6	0, 1, 3, 5, 7.5, 10
Total observations	—	—	$61 \times 10 \times 6 = 3660$

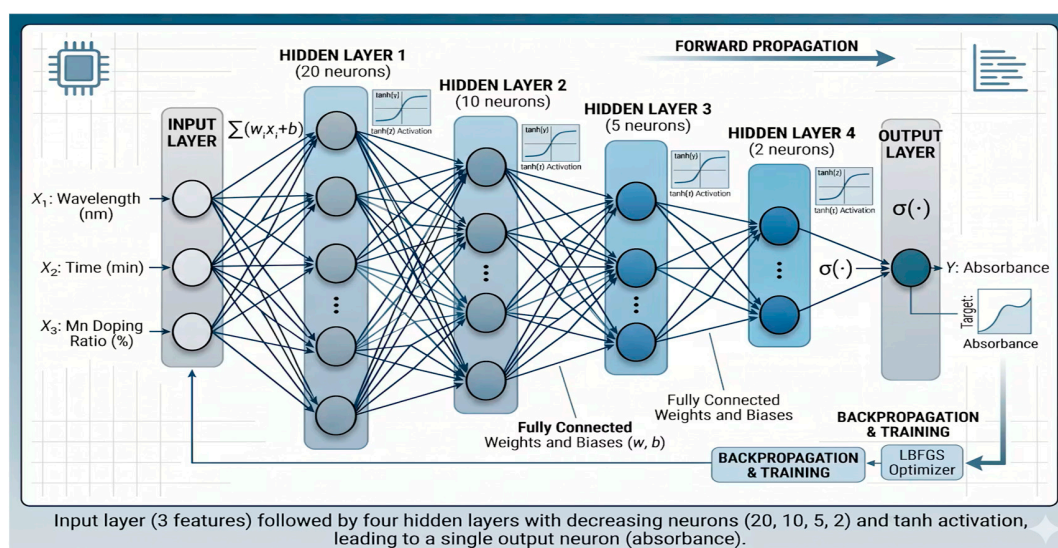


Figure 1. Structural layout of the multi-layer perceptron trained for absorbance estimation.

In the first stage, to determine the most efficient Artificial Neural Network (ANN) architecture, adhering to the *ceteris paribus* principle, each architectural parameter was modified individually, and its independent effects on model performance were observed [34]. As a result of the experiments, a four-hidden-layer structure consisting of 20, 10, 5, and 2 neurons was selected as the most suitable configuration (Figure 1). For the activation function in the hidden layers, the hyperbolic tangent ($\tanh$) function, which maps input signals to the range between -1 and 1 , was chosen [35]. The mathematical expression of the function is presented in Equation (2):

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad (2)$$

where x represents the weighted sum of the inputs to the artificial neuron.

During the model's weight optimization process, the L-BFGS (Limited-memory Broyden–Fletcher–Goldfarb–Shanno) algorithm was selected due to its ability to converge quickly on small but complex datasets. Following the training phase, the performance of the optimized MLP model was analysed using an independent test set. The model demonstrated high predictive accuracy, achieving a Coefficient of Determination (R^2) of 0.9966 and a Mean Squared Error (MSE) of 0.0030949. The MSE value is calculated using:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3)$$

Within this framework, n represents the total number of experimental data points. The terms y_i and $\hat{y}_i$ refer to the observed absorbance values and the corresponding predictions from the neural network, respectively. To evaluate the correlation between the independent variables and the dependent output, the R^2 coefficient is determined via Equation (4):

$$R^2 = 1 - \frac{\sum_{k=1}^n (y_i - \hat{y}_i)^2}{\sum_{k=1}^n (y_i - \bar{y})^2} \quad (4)$$

Here, $\bar{y}$ represents the mean of the measured absorbance values. We verified the model's ability to generalize to new data using a 5-fold cross-validation approach [36]. Across all validation folds, the R^2 scores showed a very tight standard deviation of 0.0039, suggesting the neural network architecture is structurally stable and does not suffer from overfitting.

In the final stage, to definitively prevent data leakage, the scaling and modelling processes have been integrated into a pipeline structure. Additionally, the TransformedTargetRegressor module was used to manage scale differences in the target variable (absorbance). This integrated approach ensures that no information from the test data leaks into the training phase.

In evaluating the model, the “Leave-One-Doping-Level-Out” (LODLO) cross-validation strategy—a stricter approach—was applied instead of the standard random splitting method to prevent optimism bias that could arise from spectral and temporal autocorrelation. With this method, all spectral and temporal data corresponding to a specific Mn doping level were completely isolated from the training set and used as the test set. This strategy is critical for measuring the model's out-of-sample prediction capability under experimental conditions it has never encountered before [37].

The LODLO cross-validation results have demonstrated that the MLP model possesses a high generalization capability. The model achieved a R^2 of 0.8889 and an MSE of 0.00238 on material compositions it had never seen before. These results demonstrate that the model has learned the fundamental dynamics of photocatalytic degradation rather than memorizing the data and is free from structural data leakage. This high predictive success confirms that the developed model is a reliable tool for forecasting performance under new doping levels.

2.2. Explainability of Parameters via SHAP and Permutation Importance Analyses

To move beyond the inherent “black-box” limitations of the MLP architecture, we employed a dual-validation explainability framework combining SHAP and Permutation Feature Importance. By integrating these two independent methodologies, we aim to provide an interpretable evaluation of the features driving the model's predictions. It is imperative to note that while these XAI techniques provide robust transparency into the algorithm's decision-making process, they explain the model's learned behaviour rather than establishing direct physical causality regarding the underlying photocatalytic mechanisms [38].

In the first stage, to address the inherent opacity of the MLP model, we applied SHAP analysis using the Kernel Explainer approach [39]. The goal was to break down the “black box” and quantify the exact role that UV-Vis measurement wavelength, reaction time, and Mn doping play in the results. Following the foundational Shapley principles, the impact of each feature is derived as follows (Equation (5)):

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f_{S \cup \{i\}}(x_S \cup \{i\}) - f_S(x_S)], \quad (5)$$

where ϕ_i represents the unique contribution (Shapley value) of feature i ; F serves as the total set of independent variables; S identifies any subset of features that does not include i ; and the term $f_S(x_S)$ captures the model's predicted outcome based exclusively on the features within that specific subset. The remaining terms are defined as follows: $|S|$ denotes the cardinality (number of features) of subset S ; $|F|$ denotes the total number of input features (here, $|F| = 3$); $S \cup \{i\}$ represents the union of subset S and feature i , that is, the subset S with feature i added to it; $x_{S \cup \{i\}}$ denotes the input vector containing only the features in $S \cup \{i\}$; x_S denotes the input vector containing only the features in S (excluding feature i); and the factorial expression $|S|!(|F| - |S| - 1)!$ is a weighting factor that ensures the marginal contribution of feature i is averaged fairly across all possible subsets and orderings of the full feature set. To determine the relative weight of each parameter, we aggregated the mean absolute SHAP values across the full dataset [40]. Based on the average absolute SHAP values obtained, the UV–Vis measurement wavelength emerged as the primary determinant of the model predictions with a weight of 59.6%. This was followed by the Mn addition ratio (20.5%) and reaction time (20.0%).

The nuanced relationships between the experimental parameters and the predicted absorbance are graphically elucidated in the SHAP summary plot (Figure 2). Visually reinforcing its statistical dominance, the UV–Vis measurement wavelength exhibits the widest dispersion of SHAP values, spanning broadly from approximately -0.25 to $+0.60$. For reaction time, the SHAP plot reveals a strictly monotonic negative relationship. As irradiation time increases, SHAP values drop deeply into the negative zone (reaching down to roughly -0.20), reflecting the steady, uninterrupted mineralization of the target pollutant.

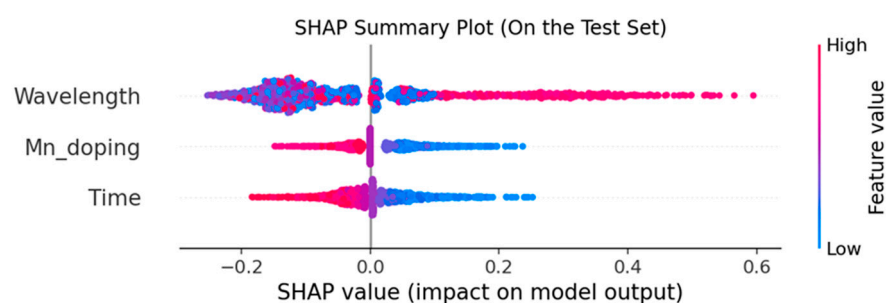


Figure 2. SHAP summary plot illustrating the influence of independent variables on the model output.

The most critical mechanistic insight from the leakage-free model emerges in the evaluation of the Mn doping ratio. The raw experimental data demonstrates that as the doping concentration increases from 0% to the optimal 7.5%, the mean absorbance decreases progressively from 0.2631 to 0.1184, confirming enhanced charge carrier trapping. At 10% Mn loading, the mean absorbance experiences a localized reversal, rising to 0.1447 due to the introduction of parasitic recombination centers. Crucially, however, the absorbance at this 10% threshold (0.1447) remains significantly lower than the undoped baseline (0.2631). The SHAP summary plot reflects this global reality with remarkable mathematical precision.

To validate these findings and eliminate the risk of algorithmic bias, a Permutation Feature Importance analysis was conducted on an independent test set [41]. This model-agnostic technique determines the global importance of a feature by measuring the increase in the model's prediction error (MSE) when the feature's values are randomly shuffled. The importance of feature j is defined as Equation (6):

$$I_j = \mathbb{E}[L(y, f(\tilde{X}^{\pi_j}))] - \mathbb{E}[L(y, f(X))], \quad (6)$$

where I_j denotes the importance score of feature j . The function L represents the loss function (in this study, Mean Squared Error, MSE), and y corresponds to the true target

values. The term $f(X)$ denotes the model predictions obtained using the original, unaltered dataset, while $f(\tilde{X}^{(\pi_j)})$ represents the predictions after randomly permuting the values of feature j , thereby disrupting its relationship with the target variable. In this framework, I_j measures the expected increase in prediction error caused by the permutation of feature j . A larger value of I_j indicates that the model's performance deteriorates substantially when this feature is perturbed, suggesting that feature j carries significant predictive information.

The permutation results strongly confirmed the SHAP hierarchy; the UV–Vis measurement wavelength received the dominant importance score of 76.7%, followed by reaction time (12.6%) and the Mn contribution ratio (10.7%). The fact that these two methods, which are based on different mathematical philosophies, agree on the fundamental results demonstrates that the established model successfully learned the physical dynamics of the process rather than simply memorizing the data points.

From a physicochemical perspective, the dominance of the UV–Vis measurement wavelength in the SHAP analysis is fully consistent with the experimental design. The spectrophotometric wavelength variable (200–800 nm) captures the full absorption profile of the MB chromophore, which varies dramatically across this range and constitutes the largest source of variance in the dataset. It must be clearly distinguished from the wavelength of the UV irradiation lamp (365 nm), which is the fixed photonic source that excites the Zn_2SnO_4 photocatalyst and initiates electron–hole pair generation. The UV lamp wavelength is not a variable in this study; the spectrophotometric scanning wavelength is. The SHAP analysis therefore reflects the dominance of the MB spectral signature as a measurement variable, not the role of UV lamp irradiation as a catalytic driver. Reaction time and Mn doping concentration act as secondary modulating factors within this framework.

In conclusion, the integration of high-accuracy MLP predictions with SHAP and permutation analyses has enabled a quantitative and transparent assessment of photocatalytic performance. This dual-validation approach ensures that the data-driven model is built on a foundation that is both statistically reliable and scientifically interpretable.

2.3. Experimental Trend Analysis of Mn-Doped Zn_2SnO_4

To further substantiate the predictive modelling framework, the experimental dataset was examined in detail to uncover the intrinsic physicochemical behaviour of Mn-doped Zn_2SnO_4 under varying operational conditions. Unlike purely statistical interpretations, this section focuses on directly observable trends and their mechanistic implications, thereby bridging the gap between data-driven modelling and semiconductor photocatalysis.

2.3.1. Analysis of Temporal Absorbance Evolution in Doped Systems

The dynamic photocatalytic response of the synthesized materials was rigorously evaluated by monitoring the UV–Vis absorbance of the methylene blue (MB) dye solution at $\lambda = 664$ nm over a 90 min irradiation interval (Figure 3a,b). It should be explicitly noted that all spectral measurements reflect the optical response of the MB dye in solution, not the intrinsic absorption of the photocatalyst itself; the decrease in absorbance at 664 nm directly corresponds to the removal of the MB chromophore from the solution phase. Figure 3a depicts the spectral decay of the pristine Zn_2SnO_4 system, which serves as a baseline for understanding the intrinsic photoactivity of the semiconductor. While a progressive reduction in the MB absorption peak is observed, the rate of degradation remains relatively moderate. In contrast, the 7.5% Mn-doped Zn_2SnO_4 specimen (Figure 3b) exhibits a markedly lower initial absorbance of approximately 0.59 at the onset of irradiation, compared to approximately 1.30 for the pristine sample. This difference does not reflect any variation in the prepared MB stock concentration, which was identical (10 mg/L) for all experiments. Rather, it is attributed to enhanced surface adsorption of MB onto the Mn-doped catalyst

surface during the 30 min dark equilibration period preceding irradiation. The incorporation of Mn ions into the Zn_2SnO_4 lattice is known to modify surface charge density and increase the density of active surface sites, which can promote electrostatic interactions with the cationic MB molecule. Consequently, a substantial fraction of MB was removed from solution via adsorption before photocatalytic irradiation commenced, resulting in the lower $t = 0$ baseline observed for the doped sample. Under these conditions, the degradation efficiency is calculated as

$$DE(\%) = \frac{A_0 - A_{90}}{A_0} \times 100, \quad (7)$$

where A_0 and A_{90} represent the absorbance values at the initial time and after 90 min of treatment, respectively. After 90 min of irradiation, the pristine sample exhibited a degradation efficiency of 61.22%, whereas the 7.5% Mn-doped sample achieved a markedly higher efficiency of 89.10%. This normalised metric eliminates the influence of differing initial absorbance values and provides a direct measure of photocatalytic activity. The substantially higher degradation efficiency of the doped sample confirms that Mn incorporation genuinely enhances the photocatalytic performance of Zn_2SnO_4 . We acknowledge that the separate quantification of adsorptive and photocatalytic contributions to the overall MB removal would provide a more complete mechanistic picture, and this will be addressed in future work.

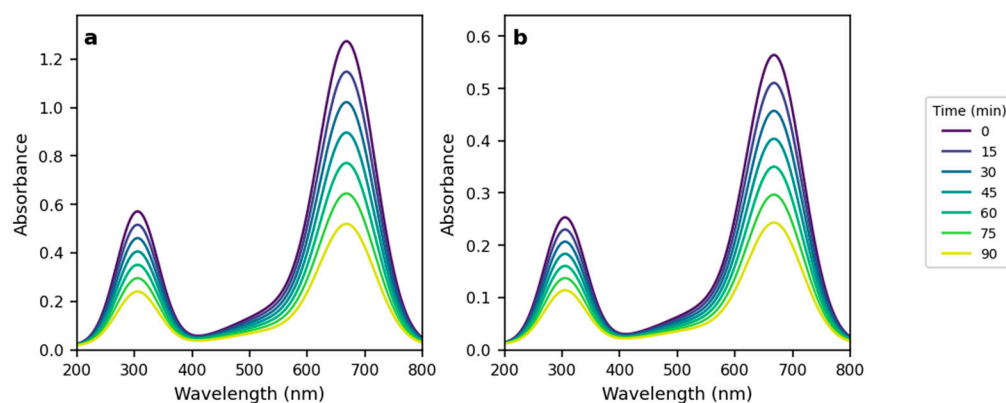


Figure 3. Temporal evolution of the UV-Vis absorption spectra during the photocatalytic degradation process for (a) the pristine Zn_2SnO_4 and (b) the optimized 7.5% Mn-doped Zn_2SnO_4 samples.

As shown in Table 2, the degradation efficiency increases progressively from 61.22% for the pristine sample to 89.10% for the 7.5% Mn-doped sample, confirming that Mn incorporation genuinely enhances photocatalytic performance. At 10% loading, efficiency decreases to 85.30%, indicating the onset of parasitic recombination. These trends are fully consistent with the SHAP and Permutation Feature Importance analyses: the UV-Vis measurement wavelength remains the dominant driver of model predictions (SHAP: 59.6%; permutation: 76.7%), while Mn doping acts as a secondary modulating factor (SHAP: 20.5%; permutation: 10.7%) that shapes the degradation response within the spectral excitation framework. From a physicochemical standpoint, Mn ions introduce intermediate energy levels within the Zn_2SnO_4 bandgap, acting as charge carrier trapping sites that suppress electron–hole recombination and extend the lifetime of reactive species. Beyond the optimal 7.5% threshold, excess Mn ions generate structural distortions that counteract these benefits. The primary role of the UV-Vis measurement wavelength—as the spectrophotometric variable that captures the full spectral variance of the MB chromophore—remains the overriding factor governing the model’s predictive output, in full agreement with the data-driven conclusions of the XAI framework. This dominance reflects a measurement property

of the dataset and is distinct from the role of the UV lamp (365 nm), which provides the fixed photonic excitation required for catalytic activity.

Table 2. Kinetic rate constants and coefficients of determination for pseudo-first-, pseudo-second-, and pseudo-third-order models applied to MB photodegradation over Mn-doped Zn₂SnO₄ samples.

Mn (%)	k_1 (min ^{−1})	R^2 (1st)	k_2 (min ^{−1})	R^2 (2nd)	k_3 (min ^{−1})	R^2 (3rd)	Deg. (%)
0	0.0105	0.9921	0.0131	0.9634	0.0326	0.8912	61.22
1	0.0124	0.9887	0.0155	0.9541	0.0384	0.8743	67.50
3	0.0158	0.9856	0.0198	0.9423	0.0490	0.8534	75.80
5	0.0195	0.9812	0.0244	0.9287	0.0605	0.8321	82.40
7.5	0.0242	0.9743	0.0302	0.9102	0.0750	0.8047	89.10
10	0.0210	0.9698	0.0263	0.9215	0.0651	0.8189	85.30

2.3.2. Structural and Optical Influence of Mn Doping

The effect of Mn incorporation on the photocatalytic performance of Zn₂SnO₄ is evaluated by comparing the degradation efficiency (%) across doping levels, calculated as $DE(\%) = \frac{A_0 - A_{90}}{A_0} \times 100$, where A_0 and A_{90} represent the absorbance values at the initial time and after 90 min of treatment, respectively, for each individual sample. This normalised metric eliminates the influence of differing initial absorbance values arising from surface adsorption during the dark equilibration period and provides a direct and unambiguous measure of photocatalytic activity under identical irradiation conditions. Figure 4 presents the UV–Vis absorption spectra recorded at $t = 90$ min for all doping levels.

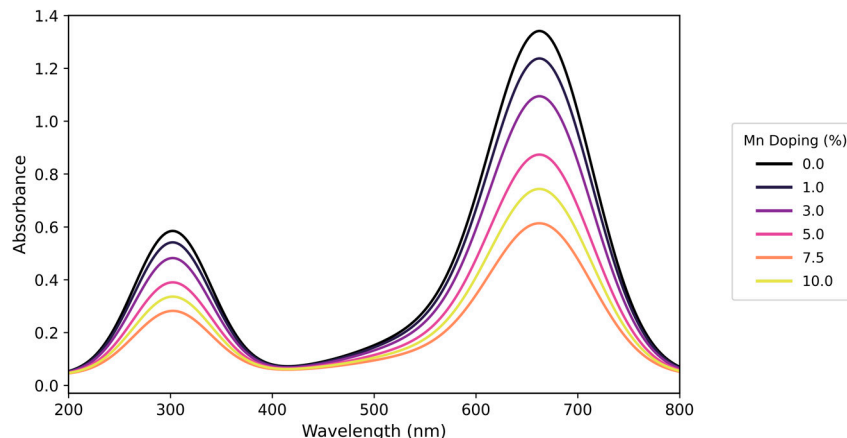


Figure 4. UV–Vis absorption spectra of Mn-doped Zn₂SnO₄ after the photocatalytic process ($t = 90$ min) as a function of various Mn doping concentrations (0% to 10%).

The degradation efficiencies derived from the kinetic analysis (Table 2) reveal a clear non-monotonic dependence on Mn loading. As the doping concentration increases from 0% to 7.5%, degradation efficiency rises progressively from 61.22% to 89.10%, representing a 1.46-fold enhancement relative to the undoped baseline. At 10% Mn loading, the degradation efficiency decreases to 85.30%, confirming that excess Mn ions act as parasitic recombination centers that partially counteract the beneficial effect of doping. This non-monotonic trend, with an optimum at 7.5% Mn, is fully consistent with the threshold behavior identified by the SHAP analysis, which showed that Mn doping acts as a secondary modulating factor (20.5% importance) within the dominant framework established by the UV–Vis measurement wavelength variable.

2.3.3. Spectral Response and Light–Material Interaction Profile

Figure 5 presents the UV–Vis absorbance spectra of the MB dye solution obtained in the presence of pristine and Mn-doped Zn_2SnO_4 samples, including the optimized 7.5% Mn-doped configuration. It should be explicitly noted that these spectra represent the optical response of the MB chromophore in solution, rather than the intrinsic absorption behavior of the photocatalyst itself. The left panel shows the full spectral range of 200–800 nm, while the right panel provides a magnified view of the UV region between 200 and 450 nm.

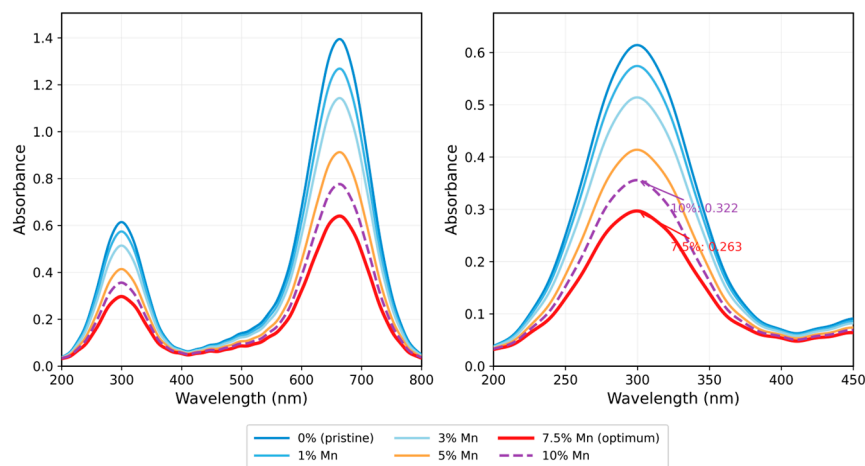


Figure 5. Representative UV–Vis absorbance spectra of the MB dye solution obtained in the presence of pristine and Mn-doped Zn_2SnO_4 samples, including the optimized 7.5% Mn-doped configuration: (left) full spectral range of 200–800 nm and (right) zoomed UV region of 200–450 nm. The optimized 7.5% Mn-doped sample exhibits the lowest absorbance, supporting its enhanced photodegradation performance.

Importantly, Figure 5 provides direct experimental support for the dominant role of the UV–Vis measurement wavelength identified by the SHAP analysis, with an importance value of 59.6%. The absorbance of the MB solution varies markedly across the wavelength axis, spanning a wide dynamic range over the measured spectrum. In contrast, the variation among different Mn doping levels at a fixed wavelength is comparatively more limited. This pronounced spectral heterogeneity is the key signal captured by the MLP model and explains why the SHAP framework assigned the highest importance to the UV–Vis measurement wavelength.

The substantial variation in MB absorbance across the 200–800 nm range indicates that different wavelength regions correspond to distinct absorption features of the MB chromophore. These spectral features generate a broad absorbance response that the MLP architecture learned to model with high accuracy. By comparison, the doping-dependent variations occur within a narrower range, which is consistent with the secondary importance assigned to Mn doping by both the SHAP analysis (20.5%) and the permutation importance analysis (10.7%). Therefore, Figure 5 serves as an experimental bridge between the spectral observations and the data-driven XAI results, providing a physically meaningful interpretation of the observed feature-importance hierarchy.

2.3.4. Evaluation of Data Distribution and Statistical Characteristics

The predictive reliability of the machine learning framework, particularly the MLP architecture employed in this research, is intrinsically linked to the underlying distribution of the experimental observations. Figure 6 below illustrates the statistical frequency and density estimation of the absorbance values recorded across the entire experimental matrix, encompassing variations in wavelength, exposure time, and Mn-doped concentrations.

The resulting histogram, overlaid with a Kernel Density Estimation (KDE) curve, reveals a distinctly right-skewed distributional profile, characterized by a high concentration of data points in the lower absorbance regimes. The specific statistical pattern is a direct manifestation of the photocatalytic degradation kinetics. By characterizing these distributional properties, we ensure that the subsequent SHAP analysis, which quantifies the 20.5% contribution of the Mn-doped ratio to the model output—notably below the uniform baseline of 33.3% expected under equal feature importance across three input variables, reflecting the dominant role of the UV–Vis measurement wavelength—is substantiated by a robust and well-distributed experimental dataset.

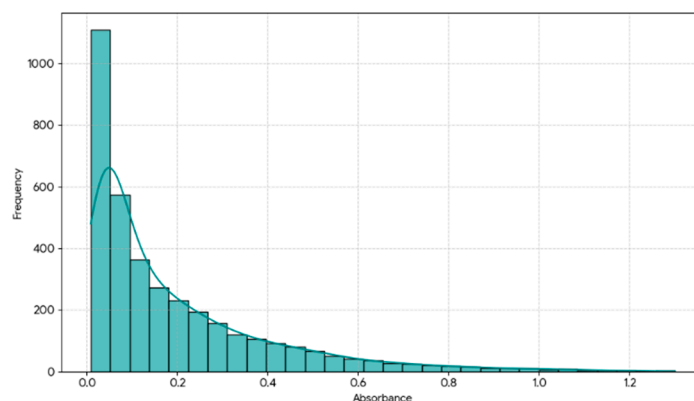


Figure 6. Statistical distribution and Kernel Density Estimation (KDE) of absorbance values across the experimental dataset.

2.3.5. Kinetic Analysis of Photocatalytic Degradation

To rigorously characterize the reaction mechanism governing the photocatalytic degradation of methylene blue over Mn-doped Zn_2SnO_4 , pseudo-first-order, pseudo-second-order, and pseudo-third-order kinetic models were applied to the time-resolved absorbance data for all Mn doping concentrations. The linearized forms of these three models are expressed as follows:

Pseudo-first-order (Langmuir–Hinshelwood):

$$\ln\left(\frac{C_0}{C}\right) = k_1 t, \quad (8)$$

$$\frac{1}{C} - \frac{1}{C_0} = k_2 t, \quad (9)$$

$$\frac{1}{C^2} - \frac{1}{C_0^2} = 2k_3 t, \quad (10)$$

where C_0 and C are the initial and time-dependent dye concentrations (proportional to absorbance under Beer–Lambert conditions), t is the irradiation time (min), and k_1 , k_2 and k_3 are the apparent rate constants for pseudo-first-, pseudo-second-, and pseudo-third-order reactions, respectively. The results of fitting these models across all doping levels are summarized in Table 2.

The kinetic data presented in Table 2 reveal that the pseudo-first-order (Langmuir–Hinshelwood) model provides the best description of the MB degradation dynamics across all Mn doping levels, consistently yielding the highest coefficients of determination (R^2 range: 0.9698–0.9921). It is important to contextualise these R^2 values relative to the MLP model performance. While the pseudo-first-order R^2 values appear high in absolute terms, they describe a single-variable linearised relationship between concentration and time at each fixed doping level, and therefore cannot capture the multivariate, nonlinear

interactions among UV–Vis measurement wavelength, reaction time, and Mn loading simultaneously. In contrast, the MLP model, evaluated under the stringent Leave-One-Doping-Level-Out (LODLO) cross-validation strategy, achieved $R^2 = 0.8889$ and $MSE = 0.00238$ across the full three-dimensional parameter space, including doping levels never encountered during training. This comparison demonstrates that the MLP framework operates at a fundamentally different level of complexity, modelling the entire experimental design space rather than individual kinetic curves, and that its generalisation performance is robust despite the more demanding evaluation protocol. The pseudo-first-order rate constant k_1 exhibits a clear non-monotonic dependence on Mn loading, rising from 0.0105 min^{-1} (pristine) to 0.0242 min^{-1} (7.5% Mn), a 2.3-fold enhancement, before decreasing to 0.0210 min^{-1} at 10% loading, confirming that excess Mn ions act as parasitic recombination centres rather than beneficial trap sites.

3. Materials and Methods

3.1. Synthesis of Mn-Doped Zn_2SnO_4

Mn-doped Zn_2SnO_4 photocatalysts were synthesized via a hydrothermal method. Zinc nitrate hexahydrate ($\text{Zn}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$, Sigma-Aldrich, St. Louis, MO, USA) and tin (IV, Sigma-Aldrich, St. Louis, MO, USA) chloride pentahydrate ($\text{SnCl}_4 \cdot 5\text{H}_2\text{O}$, Sigma-Aldrich, St. Louis, MO, USA) were used as Zn and Sn precursors, respectively, while manganese nitrate ($\text{Mn}(\text{NO}_3)_2 \cdot 4\text{H}_2\text{O}$, Sigma-Aldrich, St. Louis, MO, USA) served as the dopant source. Initially, appropriate amounts of zinc nitrate and tin chloride were dissolved in deionized water under continuous magnetic stirring until a clear solution was obtained. Subsequently, manganese nitrate was introduced into the solution to achieve the desired Mn doping levels. The mixture was stirred for 30 min to ensure homogeneous dispersion of all metal ions. The pH of the solution was adjusted to 10 using sodium hydroxide (NaOH, Sigma-Aldrich, St. Louis, MO, USA). The resulting suspension was transferred into a Teflon-lined stainless-steel autoclave and maintained at 180°C for 18 h. After the hydrothermal treatment, the autoclave was allowed to cool naturally to room temperature. The obtained precipitate was collected by centrifugation and washed several times with deionized water and ethanol to remove residual impurities. The product was then dried at 70°C for 12 h. Finally, the dried powder was calcined at 500°C for 3 h in air atmosphere to obtain crystalline Mn-doped Zn_2SnO_4 .

3.2. Photocatalytic Experimental Procedure

The photocatalytic activity of Mn-doped Zn_2SnO_4 was evaluated by monitoring the degradation of methylene blue (MB, $\text{C}_{16}\text{H}_{18}\text{ClN}_3\text{S}$, $\text{MW} = 319.85 \frac{\text{g}}{\text{mol}}$, Sigma-Aldrich, St. Louis, MO, USA) as a model organic pollutant in aqueous solution under light irradiation. MB is a cationic thiazine dye with the following structural formula: 3,7-bis(dimethylamino)phenothiazin-5-ium chloride. Its core structure consists of a phenothiazine ring system bearing two dimethylamino groups at positions 3 and 7, a sulfur atom and a nitrogen atom within the central heterocyclic ring, and a chloride counterion balancing the positive charge of the aromatic cation. In a typical experiment [42,43], 50 mg of photocatalyst was dispersed in 100 mL of MB solution at an initial concentration of 10 mg/L. The suspension pH was adjusted to 7.0 ± 0.1 using dilute HCl or NaOH solutions. Photocatalytic experiments were conducted under UV irradiation provided by a 300 W mercury lamp (emission maximum at 365 nm) positioned 15 cm above the open-top cylindrical quartz reactor. The incident light intensity at the solution surface was measured as 1.2 mW/cm^2 .

A known amount of photocatalyst was dispersed in the solution and continuously stirred at 400 rpm using a magnetic stirrer (IKA, Staufen, Germany) to ensure uniform

suspension. Before irradiation, the suspension was kept in the dark for 30 min to establish adsorption–desorption equilibrium between the catalyst surface and the pollutant molecules. The reaction system was then exposed to light irradiation. At specific time intervals, 3 mL aliquots were withdrawn, centrifuged at 5000 rpm for 5 min to remove catalyst particles, and analyzed using UV–Vis spectroscopy at the characteristic absorption wavelength of MB ($\lambda = 664$ nm). The decrease in absorbance intensity was used to quantify the degradation efficiency.

3.3. Statistical Analysis of the Experimental Parameters

Before transitioning to complex non-linear modelling, establishing the baseline linear dependencies among the experimental variables is a critical statistical step [19]. To quantify these interactions, a Pearson Correlation Analysis was conducted to evaluate how the UV–Vis measurement wavelength, reaction time, and Mn doping ratio individually influence the photodegradation efficiency (Figure 7). The findings confirmed that all selected independent parameters maintain a statistically significant linear relationship with the dependent variable ($p < 0.05$). Statistical analyses revealed a significant positive correlation between the UV–Vis measurement wavelength and photodegradation efficiency, with a correlation coefficient of 0.3311. In contrast, reaction time (-0.2723) and Mn doping concentration (-0.2307) exhibited a linear inverse relationship with photodegradation efficiency. A detailed breakdown of these correlation results and significance thresholds is presented in Table 3.

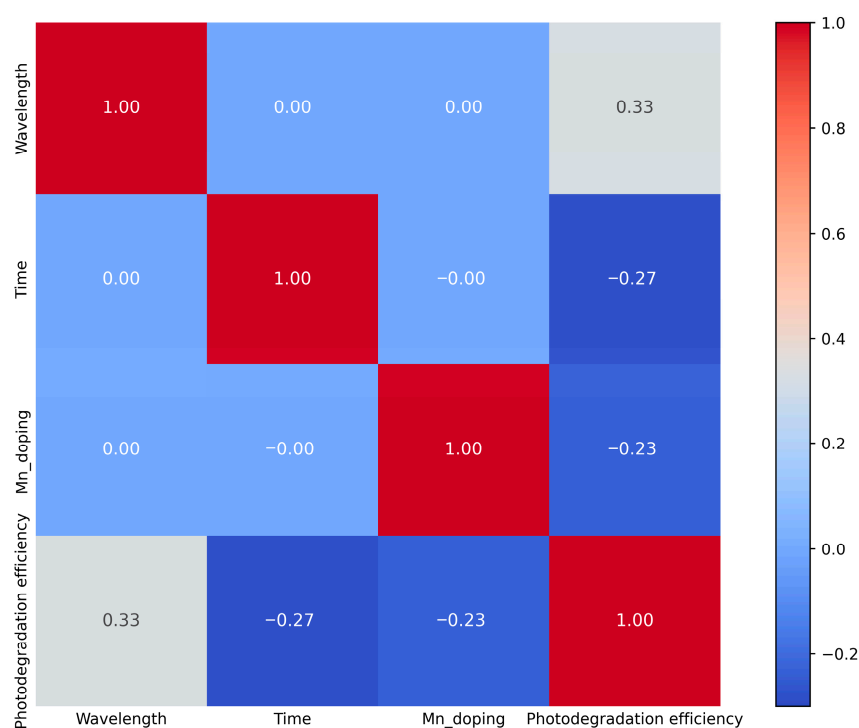


Figure 7. Correlation matrix showing the linear relationships among UV–Vis measurement wavelength, reaction time, Mn doping ratio, and photodegradation efficiency.

The F -statistic (F), used in both ANOVA and MLR analyses, is defined as the ratio of the between-group mean square ($MS_{between}$) to the within-group mean square (MS_{within}), as given:

$$F = \frac{MS_{between}}{MS_{within}} = \frac{\frac{SS_{between}}{k-1}}{\frac{MS_{within}}{N-k}}, \quad (11)$$

where SS_{between} denotes the sum of squares between groups and is calculated as:

$$SS_{\text{between}} = \sum_j n_j (\bar{y}_j - \bar{y})^2 \quad (12)$$

Here, n_j is the number of observations in group j , $\bar{y}_j$ is the mean of group j , and $\bar{y}$ is the overall (grand) mean.

Similarly, SS_{within} represents the sum of squares within groups and is defined as:

$$SS_{\text{within}} = \sum_j \sum_i (y_{ij} - \bar{y}_j)^2, \quad (13)$$

where y_{ij} is the i -th observation in group j .

In addition, k denotes the number of groups (i.e., the number of levels of the independent variable), and N is the total number of observations.

Table 3. Pearson correlation analysis showing the linear relationships and significance levels between independent experimental parameters and photodegradation efficiency.

Independent Variable	Dependent Variable	Correlation Coefficient (r)	p -Value	Significance ($p < 0.05$)
UV-Vis measurement wavelength	Photodegradation efficiency	0.3311	2.4195×10^{-94}	Statistically Significant
Time	Photodegradation efficiency	−0.2723	3.2033×10^{-63}	Statistically Significant
Mn_doping	Photodegradation efficiency	−0.2307	1.9703×10^{-45}	Statistically Significant

A large F -value indicates that the variance explained by the grouping factor is considerably greater than the unexplained, or residual, variance. This suggests that the grouping factor has a statistically significant effect on the dependent variable. The associated p -value, given in Equation (14), represents the probability of obtaining an F -statistic at least as large as the observed value under the null hypothesis that the group means are equal.

$$p = P(F_{k-1, N-k} \geq F_{\text{obs}}). \quad (14)$$

Here, F_{obs} denotes the observed F -statistic, while $F_{k-1, N-k}$ follows an F -distribution with $k - 1$ between-group degrees of freedom and $N - k$ within-group degrees of freedom. In this expression, $P(\cdot)$ denotes the probability operator, meaning “the probability that the event inside the parentheses occurs.” Therefore, the p -value is the probability that an F -distributed random variable with the corresponding degrees of freedom takes a value greater than or equal to F_{obs} , assuming that the null hypothesis is true.

Since the obtained p -value is below 0.05, the null hypothesis is rejected, indicating that the factor has a statistically significant influence on photodegradation efficiency.

ANOVA was applied to isolate the individual effects of experimental parameters on photodegradation efficiency [44]. The data presented in Table 4 demonstrate that the Mn doping ratio, reaction time, and UV-Vis measurement wavelength lead to statistically significant variations in photodegradation efficiency. The structural modification achieved through Mn doping was confirmed to have a significant effect on the catalytic optical response with a value of $F = 47.77$ ($p = 5.18 \times 10^{-48}$). Similarly, it has been mathematically determined that reaction time is a decisive factor directly affecting the process outcomes ($F = 33.09$, $p < 0.05$). The most striking deviation in the analysis was observed in the UV-Vis measurement wavelength parameter, yielding a high F -statistic of 165.20 ($p < 0.001$) [45,46].

Table 4. Summary of the analysis of variance regarding the levels of statistical significance of UV–Vis measurement wavelength, reaction time, and Mn addition ratio on photodegradation efficiency.

Source of Variation	F-Statistic	p-Value	Significance Level	Result
UV–Vis measurement wavelength	165.1957	<0.001	$p < 0.05$	Statistically Significant
Time	33.0869	1.8714×10^{-56}	$p < 0.05$	Statistically Significant
Mn Doping Ratio	47.7713	5.1812×10^{-45}	$p < 0.05$	Statistically Significant

In addition to individual variances, an MLR analysis was conducted to assess the system’s predictive capacity within a linear framework [47]. The high *F*-statistic (378.5, $p < 0.001$) produced by the model confirms that the selected independent variables (UV–Vis measurement wavelength, reaction time, and Mn ratio) are the cornerstones of the experimental design. The statistical metrics detailed in Table 5 clearly define the weights and directions of influence of these variables relative to the mathematical baseline. However, the regression model’s total explanatory power remains quite limited. The obtained R^2 value of 0.237 indicates that the independent variables can linearly account for less than one-fourth of the total experimental variance. This low value starkly highlights the extremely complex and nonlinear nature of photocatalytic interactions, necessitating the use of MLP to achieve the targeted prediction accuracy [48].

Table 5. Statistical metrics for linear regression analysis, including the weights of photocatalytic variables and their 95% confidence intervals.

Parameter	Coefficient	Standard Error	t-Statistic	p-Value	95% Confidence Interval
Intercept	0.1455	0.011	13.262	<0.001	[0.124, 0.167]
UV–Vis measurement wavelength	0.0004	1.72×10^{-5}	22.917	<0.001	[0.000, 0.000]
Time	−0.0020	0.000	−18.846	<0.001	[−0.002, −0.002]
Mn_doping	−0.0138	0.001	−15.972	<0.001	[−0.015, −0.012]

3.4. Computational Details

All statistical evaluations, machine learning modelling, and XAI analyses were conducted computationally using the Python programming language (version 3.10). Initial data manipulation, cleaning, and matrix operations were executed using the pandas and numpy libraries. The MLP architecture, as well as the Permutation Feature Importance computations, was constructed and trained utilizing the scikit-learn machine learning framework. To rigorously prevent data leakage and ensure that the test set remained completely isolated from the training distribution during cross-validation, the model was encapsulated within a scikit-learn Pipeline combined with the TransformedTargetRegressor module for dynamic data scaling. Furthermore, the mechanistic interpretation of the neural network was performed using the SHAP library, specifically deploying the KernelExplainer module to extract marginal feature contributions. All graphical representations, including the SHAP summary plots and spectral distribution curves, were generated using the matplotlib library.

4. Conclusions

This study successfully analysed the degradation performance of the Mn-doped Zn_2SnO_4 photocatalyst by integrating comprehensive statistical analyses, nonlinear machine learning modelling, and experimental findings. Initial statistical evaluations confirmed the highly nonlinear nature of the system. ANOVA results indicated that the UV–Vis measurement wavelength ($F = 165.20$), Mn loading ratio ($F = 47.77$), and reaction time

($F = 33.08$) each exerted a statistically significant influence on photodegradation efficiency ($p < 0.05$), confirming that all three parameters are meaningful determinants of the photocatalytic system's response; however, it was determined that the standard MLR model could explain only a small portion of the total variance ($R^2 = 0.237$). The optimized MLP architecture applied to overcome these linear limitations demonstrated superior performance in capturing complex degradation dynamics. The model's reliability was tested using the "Leave-One-Doping-Level-Out" (LODLO) cross-validation strategy, implemented to prevent data leakage that could arise from spectral and temporal correlations. The values of $R^2 = 0.8889$ and $MSE = 0.00238$ obtained from this rigorous evaluation demonstrate that the model can be successfully generalized to experimental conditions it has never encountered before and possesses a robust structure free from overfitting.

One of the study's most significant contributions is that it decodes the decision-making mechanism of the prediction algorithm using a dual-verified XAI framework that incorporates SHAP and Permutation Feature Importance analyses. Both methodologies identified the UV–Vis measurement wavelength as the dominant input feature driving model predictions (approximately 60% SHAP importance and a 77% permutation-based error increase). It must be emphasized that this wavelength variable refers exclusively to the spectral scanning wavelength of the UV–Vis spectrophotometer used to monitor the concentration of MB over the 200–800 nm range. This variable is distinct from the fixed wavelength of the UV lamp (365 nm), which provides the photonic excitation required to activate the Zn_2SnO_4 photocatalyst during the degradation reaction. These are two entirely distinct experimental parameters: the UV lamp provides the fixed photonic energy required to generate electron–hole pairs in the Zn_2SnO_4 lattice, while the UV–Vis spectrophotometer independently scans the absorption spectrum of the MB dye solution to quantify its concentration. The dominance of the spectrophotometric wavelength in the SHAP analysis reflects the strong spectral dependence of MB absorbance across the measurement range, which provides the largest source of variance in the dataset. Following this spectral variable, reaction time and Mn doping concentration acted as secondary modulating factors, each contributing approximately 20% to the model's predictive output, consistent with their roles in governing the kinetics and efficiency of the photocatalytic degradation process.

In conclusion, the transparent and double-verified methodology presented in this study demonstrates that machine learning can serve as a reliable tool for interpreting the complex physicochemical behaviour of doped semiconductors without remaining an opaque "black box." This developed framework provides an adaptable guide for the design of next-generation photocatalysts. Future studies are expected to further validate its practical applicability by testing this explainable architecture against real and multi-component wastewater matrices.

Author Contributions: Conceptualization, O.B., M.Z., H.A. and S.H.; Methodology, O.B. and S.K.; Software, M.Z. and S.K.; Validation, C.O.; Formal analysis, H.A.; Investigation, S.H.; Data curation, S.K.; Writing—original draft, M.Z. and S.H.; Visualization, C.O. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by Dr. Hakan Aydın.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ANOVA	One-Way Analysis of Variance
ANN	Artificial Neural Network
KDE	Kernel Density Estimation
L-BFGS	Limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm
LODLO	Leave-One-Doping-Level-Out (cross-validation strategy)
MB	Methylene Blue
ML	Machine Learning
MLP	Multi-Layer Perceptron
MLR	Multiple Linear Regression
MSE	Mean Squared Error
OFAT	One-Factor-At-a-Time
ROS	Reactive Oxygen Species
RSM	Response Surface Methodology
SHAP	SHapley Additive exPlanations
SVR	Support Vector Regression
UV	Ultraviolet
XAI	Explainable Artificial Intelligence

References

1. Qi, Z.; An, Z.; Huang, B.; Wu, M.; Wu, Q.; Jiang, D. Visible-Light-Catalyzed Synthesis of 1,3-Benzoxazines via Formal [4 + 2] Cycloaddition of Oximes with o-Hydroxybenzyl Alcohols. *Org. Biomol. Chem.* **2023**, *21*, 6419–6423. [\[CrossRef\]](#)
2. Preethi, G.; Balan, R.; Nagaswarupa, H.P. Hydrothermal Synthesis of Zn₂SnO₄/ZnO Composite for the Degradation of Organic Pollutant Methylene Blue under UV Irradiation. *Mater. Today Proc.* **2021**, *47*, 4566–4570. [\[CrossRef\]](#)
3. Nar, S.; Sahin, O.; Horoz, S. Determination optimum ni concentration in Zn₂SnO₄/Ni-doped Sb₂S₃ thin films with different Ni concentrations using incident photons to current efficiency (IPCE) and current density (J)-voltage (V) measurements. *Chalcogenide Lett.* **2018**, *15*, 491–497.
4. Fakhrzad, M.; Navidpour, A.H.; Tahari, M.; Abbasi, S. Synthesis of Zn₂SnO₄ Nanoparticles Used for Photocatalytic Purposes. *Mater. Res. Express* **2019**, *6*, 95037. [\[CrossRef\]](#)
5. Kim, T.-G.; Samuel, E.; Park, C.-W.; Joshi, B.; Kim, M.-W.; Swihart, M.T.; Yoon, S.S. Supersonically Sprayed Zn₂SnO₄/SnO₂/Carbon Nanotube Films for High-Efficiency Water Splitting Photoanodes. *J. Alloys Compd.* **2020**, *828*, 154374. [\[CrossRef\]](#)
6. Achouri, F.; Corbel, S.; Balan, L.; Mozet, K.; Girot, E.; Medjahdi, G.; Ben Said, M.; Ghrabi, A.; Schneider, R. Porous Mn-Doped ZnO Nanoparticles for Enhanced Solar and Visible Light Photocatalysis. *Mater. Des.* **2016**, *101*, 309–316. [\[CrossRef\]](#)
7. Orak, C.; Oğuz, T.; Horoz, S. Facile Synthesis of Mn-Doped CdS Nanoparticles on Carbon Quantum Dots: Towards Efficient Photocatalysis. *J. Aust. Ceram. Soc.* **2024**, *60*, 1657–1667. [\[CrossRef\]](#)
8. Horoz, S.; Dai, Q.; Maloney, F.S.; Yakami, B.; Pikal, J.M.; Zhang, X.; Wang, J.; Wang, W.; Tang, J. Absorption Induced by Mn Doping of ZnS for Improved Sensitized Quantum-Dot Solar Cells. *Phys. Rev. Appl.* **2015**, *3*, 24011. [\[CrossRef\]](#)
9. Vien, L.T.T.; Tu, N.; Viet, D.X.; Anh, D.D.; Nguyen, D.H.; Huy, P.T. Mn²⁺-Doped Zn₂SnO₄ Green Phosphor for WLED Applications. *J. Lumin.* **2020**, *227*, 117522. [\[CrossRef\]](#)
10. Sharma, S.; Pandey, H.; Kumar, M.; Chhoker, S. Room Temperature Ferromagnetism and Electrical Properties of Mn-Doped Zn₂SnO₄ Nanorods. *Superlattices Microstruct.* **2018**, *120*, 161–169. [\[CrossRef\]](#)
11. Masjedi-Arani, M.; Salavati-Niasari, M. Metal (Mn, Co, Ni and Cu) Doped ZnO-Zn₂SnO₄-SnO₂ Nanocomposites: Green Sol-Gel Synthesis, Characterization and Photocatalytic Activity. *J. Mol. Liq.* **2017**, *248*, 197–204. [\[CrossRef\]](#)
12. Junploy, P.; Phuruangrat, A.; Plubphon, N.; Thongtem, S.; Thongtem, T. Photocatalytic Degradation of Methylene Blue by Zn₂SnO₄-SnO₂ System under UV Visible Radiation. *Mater. Sci. Semicond. Process.* **2017**, *66*, 56–61. [\[CrossRef\]](#)
13. He, G.; Wen, Y.; Ma, C.; Li, X.; Gao, L.; Sun, Z. Photocatalytic Hydrogen Evolution of Nanoporous CoFe₂O₄ and NiFe₂O₄ for Water Splitting. *Int. J. Hydrogen Energy* **2021**, *46*, 5369–5377. [\[CrossRef\]](#)
14. Kutluay, S.; Horoz, S.; Şahin, Ö.; Ekin, A.; Ece, M.Ş. Highly Improved Solar Cell Efficiency of Mn-Doped Amine Groups-Functionalized Magnetic Fe₃O₄@SiO₂ Nanomaterial. *Int. J. Energy Res.* **2021**, *45*, 20176–20185. [\[CrossRef\]](#)
15. Arora, I.; Kumar, P. Enhancement of Mn²⁺ Contributions on Improvement of Electrical Characteristics for Sol–Gel Deposited Zn₂–xMnxSnO₄ Nanostructured Films. *J. Mater. Sci. Mater. Electron.* **2020**, *31*, 12725–12734. [\[CrossRef\]](#)

16. Orak, C.; Yüksel, A. Comparison of Photocatalytic Performances of Solar-Driven Hybrid Catalysts for Hydrogen Energy Evolution from 1,8-Diazabicyclo[5.4.0]Undec-7-Ene (DBU) Solution. *Int. J. Hydrogen Energy* **2022**, *47*, 8841–8857. [\[CrossRef\]](#)
17. Orak, C.; Yüksel, A. Box–Behnken Design for Hydrogen Evolution from Sugar Industry Wastewater Using Solar-Driven Hybrid Catalysts. *ACS Omega* **2022**, *7*, 42489–42498. [\[CrossRef\]](#)
18. Bakır, R.; Orak, C. Stacked Machine Learning Approach for Predicting Evolved Hydrogen from Sugar Industry Wastewater. *Int. J. Hydrogen Energy* **2024**, *85*, 75–87. [\[CrossRef\]](#)
19. Bakır, R.; Orak, C.; Horoz, S. Enhancing Photocatalytic Degradation of Hazardous Pollutants with Green-Synthesized Catalysts: A Machine Learning Approach. *J. Environ. Manag.* **2025**, *385*, 125695. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Javed, M.F.; Shahab, M.Z.; Asif, U.; Najeh, T.; Aslam, F.; Ali, M.; Khan, I. Evaluation of Machine Learning Models for Predicting TiO₂ Photocatalytic Degradation of Air Contaminants. *Sci. Rep.* **2024**, *14*, 13688. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Al Haiqi, O.; Nour, A.H.; Ayodele, B.V.; Bargaa, R. Bayesian Regularization-Trained Multi-Layer Perceptron Neural Network Predictive Modelling of Phenol Degradation Using ZnO/Fe₂O₃ Photocatalyst. *J. Phys. Conf. Ser.* **2020**, *1529*, 052058. [\[CrossRef\]](#)
22. Das, S.; Moon, S.; Kaur, R.; Sharma, G.; Kumar, P.; Lavrenčič Štangar, U. Artificial Neural Network Modeling of Photocatalytic Degradation of Pollutants: A Review of Photocatalyst, Optimum Parameters and Model Topology. *Catal. Rev.* **2025**, *67*, 544–578. [\[CrossRef\]](#)
23. Buhrmester, V.; Münch, D.; Arens, M. Analysis of Explainers of Black Box Deep Neural Networks for Computer Vision: A Survey. *Mach. Learn. Knowl. Extr.* **2021**, *3*, 966–989. [\[CrossRef\]](#)
24. Li, J.; Liu, X.; Wang, H.; Sun, Y.; Dong, F. Prediction and Interpretation of Photocatalytic NO Removal on G-C₃N₄-Based Catalysts Using Machine Learning. *Chin. Chem. Lett.* **2024**, *35*, 108596. [\[CrossRef\]](#)
25. Hosseini, S.A.; Akbari, M. ZnO/Mg-Al Layered Double Hydroxides as a Photocatalytic Bleaching of Methylene Orange—A Black Box Modeling by Artificial Neural Network. *Bull. Chem. React. Eng. Catal.* **2016**, *11*, 299–315. [\[CrossRef\]](#)
26. Armaković, S.J.; Armaković, S.; Bilić, A.; Savanović, M.M. ZnO-Based Photocatalysts: Synergistic Effects of Material Modifications and Machine Learning Optimization. *Catalysts* **2025**, *15*, 793. [\[CrossRef\]](#)
27. Kalasampath, K.; Spoorthi, K.N.; Sajeev, S.; Kuppa, S.S.; Ajay, K.; Maruthamuthu, A. A Literature Review on Applications of Explainable Artificial Intelligence (XAI). *IEEE Access* **2025**, *13*, 41111–41140. [\[CrossRef\]](#)
28. Salih, A.M.; Raisi-Estabragh, Z.; Galazzo, I.B.; Radeva, P.; Petersen, S.E.; Lekadir, K.; Menegaz, G. A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME. *Adv. Intell. Syst.* **2025**, *7*, 2400304. [\[CrossRef\]](#)
29. Vassilev-Galindo, V.; Llorca, J. Explainable Artificial Intelligence for Materials Discovery: Application to Catalysts for the HER and ORR. *Chem. Sci.* **2026**, *17*, 1058–1072. [\[CrossRef\]](#)
30. Rehman, A.; Iqbal, M.A.; Haider, M.T.; Majeed, A. Artificial Intelligence-Guided Supervised Learning Models for Photocatalysis in Wastewater Treatment. *AI* **2025**, *6*, 258. [\[CrossRef\]](#)
31. Hornik, K.; Stinchcombe, M.; White, H. Multilayer Feedforward Networks Are Universal Approximators. *Neural Networks* **1989**, *2*, 359–366. [\[CrossRef\]](#)
32. Koukaras, P.; Tjortjis, C. Data Preprocessing and Feature Engineering for Data Mining: Techniques, Tools, and Best Practices. *AI* **2025**, *6*, 257. [\[CrossRef\]](#)
33. Karateke, S.; Zontul, M.; Bozkurt, N.E.; Aslan, Z. Wavelet-ANFIS Hybrid Model for MODIS NDVI Prediction. *J. Appl. Remote Sens.* **2021**, *15*, 024519. [\[CrossRef\]](#)
34. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
35. Nwankpa, C.; Ijomah, W.; Gachagan, A.; Marshall, S. Activation Functions: Comparison of Trends in Practice and Research for Deep Learning. *arXiv* **2018**, arXiv:1811.03378. [\[CrossRef\]](#)
36. Kohavi, R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *Ijcai* **1995**, *14*, 1137–1145.
37. Jiang, D.; Yang, C.; Man, X.; Li, X.; Wang, H.; Song, Y.; Qi, Z. Fe(NO₃)₃·9H₂O-mediated visible-light photocatalysis for the nitration of cyclobutanols: Access to functionalized nitrocyclobutenes. *Org. Biomol. Chem.* **2025**, *24*, 152–156. [\[CrossRef\]](#)
38. Cravero, F.; Ponzoni, I.; Diaz, M.F.; Vazquez, G.E. From Prediction to Understanding: A Review of XAI Applications and Innovations in Materials Science. *Comput. Mater. Sci.* **2026**, *267*, 114493. [\[CrossRef\]](#)
39. Lundberg, S.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems*; Neural Information Processing Systems Foundation, Inc.: La Jolla, CA, USA, 2017.
40. Molnar, C. *Interpretable Machine Learning A Guide for Making Black Box Models Explainable*; Lulu.com: Durham, NC, USA, 2022.
41. Mehdiyev, N.; Majlatow, M.; Fettke, P. Integrating Permutation Feature Importance with Conformal Prediction for Robust Explainable Artificial Intelligence in Predictive Process Monitoring. *Eng. Appl. Artif. Intell.* **2025**, *149*, 110363. [\[CrossRef\]](#)
42. Chen, S.; Deng, J.; Wang, L.; Ma, X.; Xu, M.; Xu, D.; Wang, H.; Shi, W. Comparison of Vacuum Ultraviolet-Based Oxidation as Ultrafiltration Pretreatment in Real Water: Performance, Risk Assessment, and Mechanistic Insights. *Water Res.* **2026**, *298*, 125801. [\[CrossRef\]](#)
43. Choi, W.; Termin, A.; Hoffmann, M.R. The Role of Metal Ion Dopants in Quantum-Sized TiO₂: Correlation between Photoreactivity and Charge Carrier Recombination Dynamics. *J. Phys. Chem.* **1994**, *98*, 13669–13679. [\[CrossRef\]](#)

44. Baytar, O.; Şahin, Ö.; Horoz, S.; Kutluay, S. High-Performance Gas-Phase Adsorption of Benzene and Toluene on Activated Carbon: Response Surface Optimization, Reusability, Equilibrium, Kinetic, and Competitive Adsorption Studies. *Environ. Sci. Pollut. Res.* **2020**, *27*, 26191–26210. [[CrossRef](#)] [[PubMed](#)]
45. Naresh Yadav, D.; Anand Kishore, K.; Saroj, D. A Study on Removal of Methylene Blue Dye by Photo Catalysis Integrated with Nanofiltration Using Statistical and Experimental Approaches. *Environ. Technol.* **2021**, *42*, 2968–2981. [[CrossRef](#)] [[PubMed](#)]
46. Devi, L.G.; Kavitha, R. A Review on Non Metal Ion Doped Titania for the Photocatalytic Degradation of Organic Pollutants under UV/Solar Light: Role of Photogenerated Charge Carrier Dynamics in Enhancing the Activity. *Appl. Catal. B* **2013**, *140–141*, 559–587. [[CrossRef](#)]
47. Bezerra, M.A.; Santelli, R.E.; Oliveira, E.P.; Villar, L.S.; Escaleira, L.A. Response Surface Methodology (RSM) as a Tool for Optimization in Analytical Chemistry. *Talanta* **2008**, *76*, 965–977. [[CrossRef](#)]
48. Oh, H.-S. Introduction to Linear Regression Analysis, 5th Edition by MONTGOMERY, DOUGLAS C., PECK, ELIZABETH A., and VINING, G. GEOFFREY. *Biometrics* **2013**, *69*, 1087. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.