Article

Design of a Decision Support System to Operate a NO$_2$ Gas Sensor Using Machine Learning, Sensitive Analysis and Conceptual Control Process Modelling

Mohammad Gheibi [1] , Hadi Taghavian [2] , Reza Moezzi [2,*] , Stanislaw Waclawek [2] , Jindrich Cyrus [2] , Anna Dawiec-Lisniewska [3], Jan Koci [2] and Masoud Khaleghiabbasabadi [2]

[1] Association of Talent under Liberty in Technology (TULTECH), 10615 Tallinn, Estonia
[2] Institute for Nanomaterials, Advanced Technologies and Innovation, Technical University of Liberec, 46117 Liberec, Czech Republic
[3] Department of Advanced Materials Technologies, Faculty of Chemistry, Wroclaw University of Science and Technology, 50-373 Wroclaw, Poland
[*] Correspondence: reza.moezzi@tul.cz

Abstract: The most advantageous method for detecting dangerous gases and reducing the risk of potential environmental toxicity effects is the use of innovative gas sensing systems. However, designing effective sensors requires a complex process of synthesizing functional nanoparticles, which is a costly process. Additionally, practical operation of the toxic gas sensors always carries a significant cost along with a considerable risk of hazardous gas emissions. Machine learning algorithms may be used to accurately automate the behavior of the sensors to eliminate the abovementioned deficiencies. In the present research, there are three different factors involved in the optimization of NO$_2$ sensing by means of the response surface methodology (RSM). Two main functions of sensor efficiency, namely sensitivity and response time, are predicted according to the Fe$_3$O$_4$ additive (%), input NO$_2$ (ppm), and response time/sensitivity, and moreover, the execution of a controlling system of the sensor network using the Jacobson model is proposed. The machine learning computations are implemented by Meta.RegressionByDiscretization, M5.Rules, Lazy KStar, and Gaussian Processes algorithms. The outcomes illustrate that the best gas sensor efficiency predictions are related to M5.Rules and Lazy KStar, with a correlation coefficient of more than 96%. The best performance of machine learning computations can be found in the range of 8–10-fold in training and testing arrangements. Meanwhile, the ANOVA assessment confirmed that the most important features in the prediction of response time and sensitivity are NO$_2$ concentration and response time, respectively, with the lowest *p*-value recorded. The outcomes illustrated that with combinations of RSM, machine learning, and the Jacobson model as a controller, a decision support system can be presented for the NO$_2$ gas sensor system.

Keywords: NO$_2$; gas sensor; machine learning; historical data analysis; sensitive analysis; conceptual modelling

1. Introduction

Adverse air quality has caused enormous difficulties in human health, and unfortunately about 7 million mortalities every year [1]. Lung cancer, heart disease, respiratory infection, irritating asthma, and neurological misfunctioning are some examples of diseases that are caused by air pollution [2–4]. This pollution, which is the most harmful to children, the elderly [5,6], and people with pre-existing health issues [7], can also affect people of all ages and health conditions. Apart from the obvious health repercussions, air pollution can cause other results such as damaging agricultural crops, plants, and construction materials as well as impairing visibility in urban and peri-urban regions [8–10].

The case pollutant in this paper is NO_2, a greenhouse gas and an essential indicator of air contamination. It also shows the presence of a number of hazardous secondary air pollutants, including ozone and particulate matter such as $PM_{2.5}$ and PM_{10}. Due to all of the aforementioned reasons, networks of sensors have been utilized to monitor environmental variables [11,12].

In industrial and residential contexts, chemical or physical mode sensors are used to carry out a range of functions, including prediction, prognostics, forecasting, remaining effective life estimates, and trend analysis [13–15]. Gas-caused signals, temperature, pressure, strain, and other characteristics are integrated by the chemical and physical modes. In terms of big data, AI (artificial intelligence), ML (machine learning), and IOT (Internet of Things), the fourth industrial revolution has given rise to a paradigm shift in interpretation and data acquisition. AI, particularly machine learning, plays a key role in modelling, which leads to decision-making and prediction.

Government and environmental organizations have begun to establish air monitoring sites for quality assurance in several regions to collect air quality data. With the aid of precise instrumentation, these monitoring stations can regularly monitor air quality in the environment, assess pollutant concentrations, and deliver data to the public as a reference [16,17]. To increase the spatial density of air quality statistics, a variety of proposals have been made, including employing mobile laboratories to detect short-term air quality campaigns [18], mathematical models to interpolate data [19], and novel low-cost sensors to detect air quality [20,21]. Several scientists have concentrated on creating novel sensing materials to enhance sensor detection. It has been shown that SnO_2 nanosurfaces chemically altered by metal oxides are highly efficient materials for sensing NO_2 gas molecules [22]. In addition, Gholami et al. utilized ZnO nanoparticles to functionalize multiwall carbon nanotubes for sensing NO_x gas [23]. Although more precise gas sensing may be achieved with novel sensing materials, not all practical needs have been fully met by low-cost sensors. Contrarily, we see that AI technology has made substantial prior advancements and is effectively used in a number of industries. As a consequence, gas sensors have been combined with AI technology, particularly in terms of machine learning, to provide more precise prediction and detection. As an example, an ANN (artificial neural network) model was applied with a combination sensor consisting of four QCM (quartz crystal microbalance) sensors in [24] to detect a variety of organic pollutants. A fabricated ZnO-based sensor and a model of ANN were used in another investigation to detect the presence of H_2, CH_4, and CO gas concentrations [25]. ANN models were also used with sensors, accounting for pressure, temperature, and humidity. ANN was determined to be the most efficient technique among multiple linear regressions [26] for the tuning of affordable industrial sensors for sensing NO, CO_2, and CO molecules [27]. In another study, a ML tuning model was used to enhance sensor efficiency for affordable air quality surveillance for random forests [28]. A RNN (recurrent neural network) is an alternative deep learning technique that is widely used to resolve classification issues because of its gated-unit architecture [29–33]. Another novel method for addressing the deep learning bottleneck is ensemble modelling [34,35]. The following may be a summary of further instances in the field: Li et al. [36] estimated ground NO_2 concentration levels based on simulations of climatic factors and surface mass concentrations of nitric acid performed using the GEOSChem software (https://doi.org/10.5281/zenodo.3507501, accessed on 15 December 2022). They inputted the raster layers into a geologically and chronologically weighted comprehensive regression NN together with data from NDVI from Terra, the Sentinel-5P, Aqua, and a digital elevation model. As an alternative, Ghahremanloo et al. [37] employed deep learning methods from distant sensing information to estimate the ground daily level of NO_2 concentrations using the WRF simulated meteorological parameters via Aura satellite NO_2 retrievals, and interpolated population data. The use of these models for directly mimicking the behavior of target variables has also been considered in recent works. GEOS-Chem was used to test in Beloconi and Vounatsou's study [38] for daily NO_2 estimation as well. This model used a variety of other predictors, such as tree cover density,

nighttime light, land cover, terrain elevation, data from roads, NDVI, and climatic data. On the other hand, DSS (decision support system) has shown a deep structure to aid the operation, management, and planning stages of a process in terms of information systems that support organizational and structural decision-making activities [39,40]. The current study on gas sensing using machine learning computations with mentioned techniques, to the best of the authors' knowledge, for the first time, aims to estimate and predict the sensor performance based on sensitivity, response time and the Fe_3O_4 [41] additive concentration which can potentially increase gas sensing effectiveness. Figure S1 shows the scientometry analysis of machine learning applications in gas sensors in contemporary situations, indicating their importance in current studies in this field. Furthermore, Figure S2 indicates the contribution of countries in the field of artificial intelligence in gas-sensor designing.

This paper sequentially investigates: (i) the fabrication and implementation of nanosensors for the detection of NO_2 gas; (ii) the design of experiments and optimization followed by the study of the impact of each effective variable on the nano-gas-sensing system via the response surface methodology (RSM), and (iii) exploiting the machine learning computations using Meta.RegressionByDiscretization, M5.Rules, Lazy KStar, and Gaussian Processes algorithms. Soft computation and the system control of gas-sensor operation by means of the Jacobson model are proposed. Finally, future studies and conclusion are presented.

2. Materials and Methods

2.1. Mechanism of Sensor Creation

This paper considered the operation of a NO_2 gas sensor (Fe_3O_4-rGOQD-SO_3H) using sensitivity analysis, machine learning calculations, and conceptual modelling of the control process. The NO_2 gas sensor was fabricated in five steps. First, GO was fabricated using the Hummer method [42]. Then, functionalization, generation of some dots on the surface of GO, and magnetization of GO were performed according to our previous research [43–45]. Then, the chemical structure of the nanosensor was characterized by means of different analysis techniques such as FESEM (Figure S3), EDX, FT-IR, XRD, TGA, and HRTEM. Finally, the performance of the fabricated NO_2 gas sensor was evaluated with different percentages of Fe_3O_4 loading and considered to detect NO_2. Moreover, the NO_2 concentration for the fabricated NO_2 gas sensor was selected between 2.5 and 50 ppm. The synthetic NO_2 gas sensor was investigated as a highly effective and reusable nanosensor with different weight contents of Fe_3O_4 (99, 95, 90, and 85 wt%). The sensor evaluation results showed that the best results could be obtained by using the mass ratio (X = 15:85 wt%) of Fe_3O_4 and rGOQDs-SO_3H. Based on our results, the optimum temperature, detection range, and reactivity are improved compared to other studies. Likewise, the fabricated sensor was able to spontaneously return to its primary conditions by flowing N_2 without thermal support or chemical treatment [45]. Figure 1 shows the simultaneous efficacy of NO_2 concentration, Fe_3O_4 additive percentage, and sensitivity or response time as independent variables for gas performance estimation.
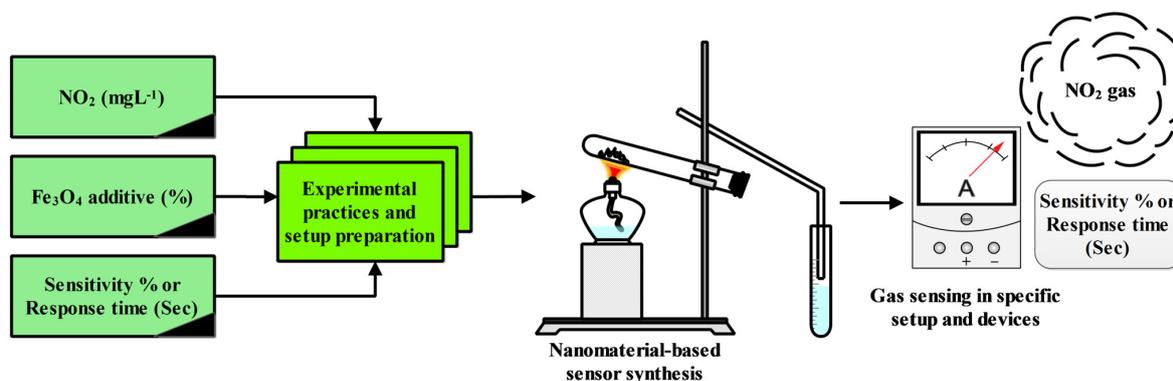
Figure 1. The effect of three factors on each other in the performance of a NO_2 gas sensor.

2.2. Optimization and Sensitive Analysis

Design Expert (version 7.0.0) was utilized for sensitivity analysis and optimization of the operation of the NO_2 gas-sensing system. In this investigation, the response surface methodology (RSM) was exploited to determine the decision support system (DSS) based on experimental values derived from the actual NO_2 nano gas-sensor in operation. The impacts of three operational factors on the performance of the gas-sensor, which were analyzed by RSM, were: (i) the percentage of Fe_3O_4 nanomagnetic agent; (ii) the concentration of NO_2 gas (mg/L); and (iii) either a percentage of sensitivity (if the response time is the answer) or seconds of response time (if the sensitivity is the answer). Moreover, RSM facilitated the optimized gas sensing process by analyzing the relationships between these three independent variables. The historical data analysis of the RSM technique was employed to find the optimum condition. Then, the historical data analysis was directed by RSM to gain optimal values with higher desirability and obtain the mathematical predictive models from ANOVA tables. Eventually, after obtaining the optimum values, machine learning computations, sensitive analysis, and conceptual modelling were executed to obtain the DSS of the NO_2 gas sensor operation.

In the last section of RSM evaluations, the optimal suggestions of the designed model were presented as per the desirability function. The function determined the accuracy of the model as per some mathematical computations. When Y_i is the predicted response as per t_i ($l_i < t_i < u_i$), desirability of Y_i is called $d_i(Y_i)$, and it is computed as Equation (1). In this equation, *s* and *b* are related to power of the equations [46].

$$d_i(Y_i) = \begin{cases} 0 & Y_i < l_i \\ \left(\frac{Y_i - l_i}{t_i - l_i}\right)^s & l_i \le Y_i \le t_i \\ \left(\frac{Y_i - u_i}{t_i - u_i}\right)^b & t_i \le Y_i \le u_i \\ 0 & u_i < Y \end{cases} \qquad (1)$$

In the following, the overall desirability of n responses (D) is computed as the geometric average (Equation (2)).

$$D = \left(\prod_{i=1}^{n} d_i(Y_i)\right)^{\frac{1}{n}} \qquad (2)$$

2.3. Machine Learning Computations

In the present study, due to the implementation of a smart system for the prediction of sensitivity and response time, some different machine learning algorithms were utilized with the application of WEKA 3.9 software [47]. In this process, first the arranged data were trained, and in the next step, the outputs were tested for the determination of the algorithms' performances. In this research, four algorithms including Meta.RegressionByDiscretization, M5.Rules, Lazy KStar, and Gaussian Processes were utilized. For machine learning computations and analysis, the data were divided into two sections, including training data and testing data. The data for both categories were provided based on practical experimental results. The application of training data, the simultaneous efficacy of variables (Figure 1), and predictive functions were modeled and the performance of each predictor algorithm was evaluated by a testing process. In the testing process, the predicted and actual values were compared through the application of statistical indicators. During both the training and testing processes, the experimental data were sorted based on the variables (input data: NO_2, Fe_3O_4 additive %, and sensitivity/response time, output: sensitivity/response time) according to Figure 1 and then machine learning algorithms were employed for the implementation of the smart soft-sensor framework.

Data discretization techniques such as Meta.RegressionByDiscretization were used as one of their classification and reduction solutions. In these data mining methods, tagging processes can be used to replace real data values. In the mentioned method, in addition to reducing the size of the data in the learning process, the attributes can also be summarized

and create better conditions for analysis, which is rarely found in other techniques [48]. In the setting of the WEKA software for this algorithm, some different machine learning features are adjusted. For example, the random seed for XVal/% Split is set to 1 and also, the number of leaves and size of the tree are controlled at 7 and 13, respectively.

M5P is known as a machine learning method based on decision trees with regression functions. In this technique, the logic of data classification is based on the priority of cause-and-effect relationships between them, which is realized in a binary environment. In this method, the data clustering structure is modeled as a tree. In the declared algorithm, standard deviation reduction (SDR) is the main criterion for the assessment of different conditions' performance (SDR = $\mathrm{sd}(M) - \sum_{M}^{M_i} \mathrm{sd}(M)$, M: Node example set and M_i: ith input subset) [49]. In this study, the classification process of the M5P model was completed using 3 rules by means of WEKA software.

The Gaussian regression method works based on the parallel solution of multi-device equations and sequential substitutions. In this method, the distribution of data is normalized, and the forecasting processes are based on the equations of this distribution [50]. Likewise, for data pattern analysis in the GP technique, kernel machines are applied.

Lazy learning is based on the generalization of search data with a deferred approach. The logic used in these algorithms is exactly the opposite of enthusiastic techniques. The term K star is applied by entropic measurement as the instance-based probability transferring parameter [51]. In the present investigation (in WEKA), -B 20 -M a is applied as the K star option and also XVal/% Split is set on 1.

2.4. Controlling System of Sensor Network

The controlling system of gas detection was designed by the application of the Jacobson model as a use-case technique. In this step of the research, the model was created using EdrawMax 6.8 software, and the main purpose of the scheme was related to controlling the NO_2 gas sensor's optimum performance. In the following, the structure of the model is illustrated as per Figure 2. As per the scheme, it can be seen that in the first step, the optimal conditions of input data include the percentage of Fe_3O_4 additive as an additive and the concentration of NO_2 as an air pollutant, while sensitivity/response time are determined based on RSM outcomes. Then, the best amounts of the features are prioritized (as per ANOVA analysis) and controlled in parallel loops in the concept of the Jacobson technique. In the Jacobson technique, the relationships among roles, goals, and actors are modeled as an action plan of an operational process (Jacobson et al., 2016).
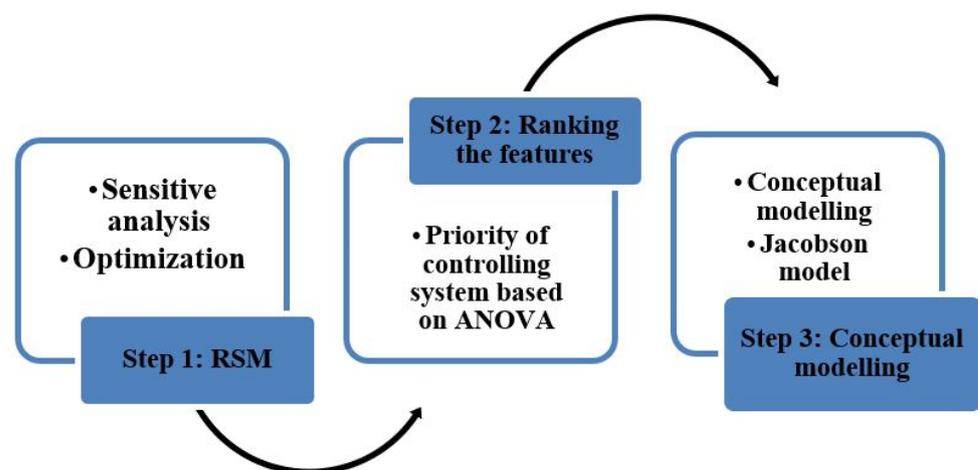
Figure 2. The structure of conceptual modelling by means of the Jacobson technique in the present research.

3. Results and Discussions

3.1. Sensitive Analysis and Optimization

The RSM analyses based on the central composite design (CCD) technique were executed on the practical results of the NO_2 nano gas-sensor. Tables 1 and 2 show the suggested parameters of the system based on the independent variables of Fe_3O_4 nanomagnetic agent (%), NO_2 gas (mg/L), and also sensitivity (%) (response time is the answer) or response time (sensitivity is the answer), respectively. The regression fitness analyses based on the linear, 2FI, quadratic, and cubic models are shown in Table 1 (for response time answer) and Table 2 (for sensitivity answer). The outcomes of descriptive statistical analysis including skewness, kurtosis, standard deviation, etc., for both response time and sensitivity are demonstrated in Table S1.

Table 1. Statistical parameters based on different regression fit models and R-factors for response time (s).

Source	Std. Dev.	R-Squared	Adjusted R-Squared	Predicted R-Squared	PRESS	
Linear	68.95459	0.870631	0.85215	0.799389	154,836	
2FI	39.13109	0.964289	0.952386	0.911022	68,675.3	
Quadratic	29.09954	0.983543	0.973669	0.78265	167,755.1	Suggested
Cubic	23.78363	0.996336	0.982411	−0.32398	1,021,879	

Table 2. Statistical parameters based on different regression fit models and R-factors for sensitivity (%).

Source	Std. Dev.	R-Squared	Adjusted R-Squared	Predicted R-Squared	PRESS	
Linear	24.05909	0.716192	0.675649	0.534586	19,933.96	
2FI	7.598312	0.975737	0.967649	0.915811	3605.856	
Quadratic	7.622766	0.97965	0.96744	0.875495	5332.603	
Cubic	2.11794	0.999476	0.997486	0.769755	9861.54	Suggested

The significance of the suggested models was statistically evaluated by the regression coefficients and standard errors of the regression models. The multiple coefficients of determination (R-squared), adjusted coefficient of determination (adjusted R-squared), and predictive R-squared are defined to check the standard error, accuracy of the models, and predictive desirability of the models, respectively.

Comparison between the R-squared, adjusted R-squared, and also predicted R-squared of different regression models reveals that the quadratic model with 0.98, 0.97, and 0.78 values (Table 1) and the cubic model with 0.999, 0.997, and 0.769 values (Table 2 were the best regression fits comparatively. Accordingly, the quadratic and cubic models were applied to acquire the mathematical equations for determining the response time and sensitivity based on Equations (3) and (4), respectively.

$$\text{Response time} = 669.31687 - 13.44598 * A - 13.88277 * B - 4.98532 * CS + 0.12005 * A * B + 0.052133 * A * CS - 0.012850 * B * CS - 0.042876 * A^2 + 0.17070 * B^2 + 0.022389 * CS^2 \tag{3}$$

$$\text{Sensitivity} = 878.60286 - 76.87784 * A - 29.19772 * B - 4.72585 * CR + 2.61581 * A * B + 0.25920 * A * CR + 0.13863 * B * CR + 2.30169 * A^2 + 0.019486 * B^2 + 8.13204 \times 10^{-3} * CR^2 - 5.26438 \times 10^{-3} * A * B * CR - 0.043170 * A^2 * B - 3.72342 \times 10^{-3} * A^2 * CR - 0.012299 * A * B^2 - 2.09463 \times 10^{-4} * A * CR^2 - 6.23848 \times 10^{-4} * B^2 * CR - 1.39499 \times 10^{-4} * B * CR^2 - 0.022116 * A^3 + 2.98711 \times 10^{-3} * B^3 - 4.53829 \times 10^{-6} * CR^3 \tag{4}$$

where A is the percentage of Fe_3O_4 nanomagnetic agents in the nano gas-sensor, B is the concentration of NO_2 gas, CS is the sensitivity percentage, and CR is the response time. Although both resultant equations for determining the response time and sensitivity showed ample R-squared and adjusted R-squared values, the low values of the Predicted

R-squared of 0.782 and 0.769 (Tables 1 and 2) decrease the trustfulness of the equations for the prediction of these functional factors in gas sensing operation. Therefore, taking advantage of AI via machine learning technology can be an applicable approach to increase the accuracy of the predictive models for the DSS.

Analysis of variance (ANOVA) results of both suggested models, (i) quadratic (for response time as the answer) and (ii) cubic (for sensitivity as the answer), are shown in Tables S1 and S2, respectively. The probability value (*p*-value) lower than 0.0001 is attributed to the significance of the model (Tables S2 and S3). None of the *p*-values of the variables in both models are <0.0001, which means that none of the variables individually have a salient impact on the system response separately. Nonetheless, a comparison between the F values of Fe_3O_4 (%), NO_2 (mg/L), and sensitivity (%) in Table S2 (0.569, 1.168, and 0.286) implies the more substantial influence of gas concentration on the response time of the gas sensing system. Likewise, the slightly higher F value of the response time than gas and magnetic agent concentration in Table S3 (6.16, 5.11, and 3.875) demonstrated the relatively higher effectiveness of the response time variable on the system's sensitivity.

The effectiveness of the system's response depends on the combinational efficacies of all integrated variables. The closeness of the values among independent factors on system response (in both regression models) provides evidence for this. The efficacy of each parameter on the system's response cannot be investigated individually. As a result, the 3D diagrams can provide a more accurate analysis of the mutual effects of the variables on the system's response.

Figure 3 indicates the system's response (sensitivity and response time) around the normal diagrams. The closeness of the experimental values to the normal diagrams of (a) and (b) signals the validity of the proposed technique. According to Figure 3b, it can be seen that in the condition of predicting the sensitivity of the sensor, the distribution of the results of the tests is not statistically absolutely normal and has some skewness. This outcome is completely based on statistical analysis, and due to the non-normality of this distribution of results, the importance of using artificial intelligence to predict these functions is more evident.
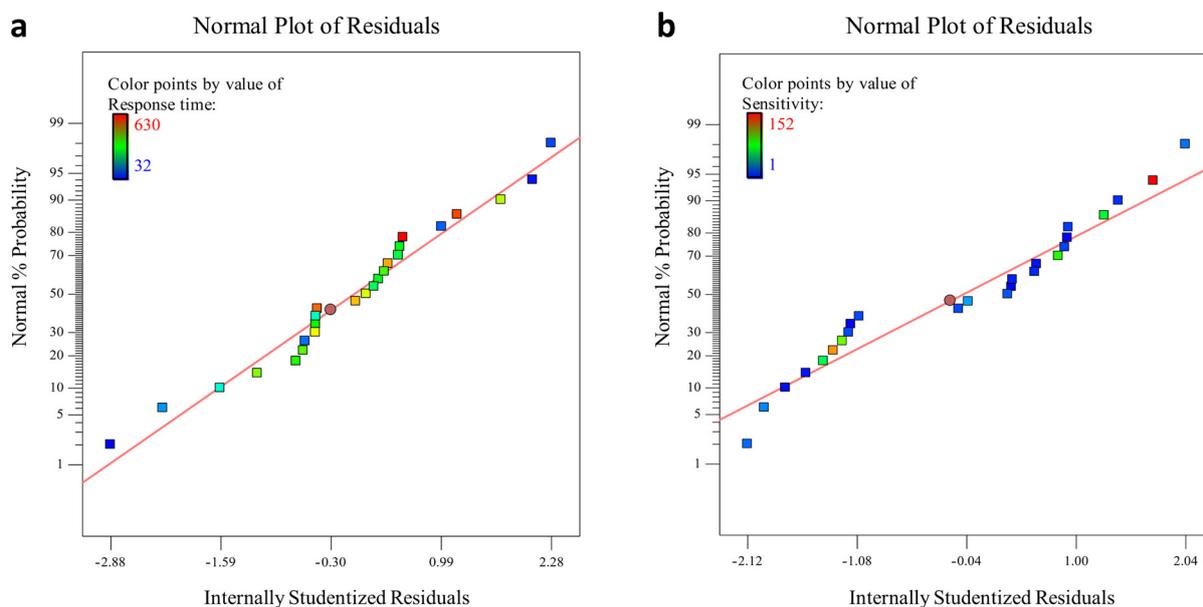
Figure 3. System's response around the normal diagrams: (**a**) Response Time and (**b**) Sensitivity.

Figure 4 represents the 3D plot of interactions among dual independent parameters on the response time and sensitivity as responses of the system, respectively. As shown in Figure 4a–c, increasing the NO_2 gas concentration leads to the mitigation of demanding time for the gas sensor system to respond. In addition, in a similar effect, the increment

of the Fe_3O_4 magnetization agent of the gas sensor contributes to decreasing the system's response time. In the same way, decreasing the sensor's sensitivity causes an increase in the time for the system to respond. Moreover, analyses of the 3D diagrams in Figure 4a–c revealed that the sharper slopes of the diagram on the side of gas concentration (Figure 4a) and sensitivity (Figure 4b) are due to their higher impact on the response time of the system. Subsequently, the effective mutual impact of both gas concentration and sensitivity variables brought a double impact on the response time diagram (Figure 4c).
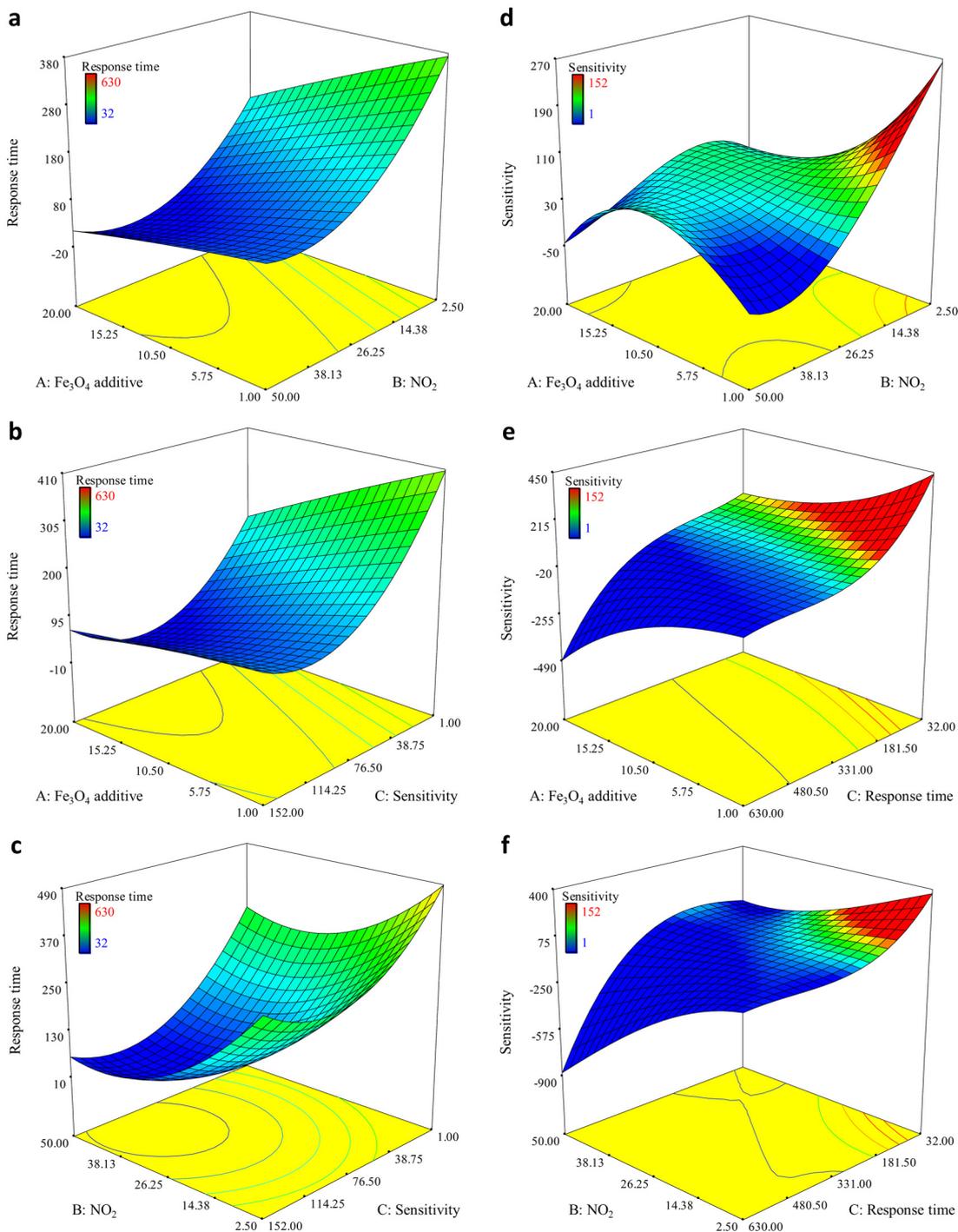
Figure 4. Sensitive analysis of the input data in the optimization model based on response time (**a**–**c**), and the outputs of dual comparison of the most significant features according to sensitivity (**d**–**f**).

Similarly, the dependency of the sensitivity of the system on the variables can be traced by pursuing the slope variation after differing the NO_2 concentration and Fe_3O_4 (Figure 4d), response time and Fe_3O_4 percentage (Figure 4e), and response time and NO_2 concentration (Figure 4f). In agreement with the F value results of Table S3, the response time shows the highest efficacy on the system's sensitivity. As shown in Figure 4d–f, a shorter response time equals a higher sensitivity of the gas sensing system. Furthermore, decreasing the NO_2 concentration increases the sensor's sensitivity for gas detection.

Tables 3 and 4 present the six calculated optimum values on the basis of the historical data analysis method of the RSM technique [52–54] for the nano gas-sensor system distinguished by response time and sensitivity, respectively. According to Table 3, the first suggested optimum values of the Fe_3O_4 agent, NO_2 gas concentration, and sensitivity are 19.44%, 25.46 mg. L^{-1}, and 95.37%, respectively, which leads to an estimated response time of 1.976 s. The desirability of this predicted response based on the suggested optimum variables is depicted in Figure 5a. The high desirability of this model (close to the value 1), particularly with a high NO_2 gas concentration and amount of Fe_3O_4 agent and constant sensitivity (95.37), implies the model's accuracy.

Table 3. Suggested optimized values for desirable lower response time of the gas sensor system.

Number	Fe$_3$O$_4$ Additive	NO$_2$	Sensitivity	Response Time
1	19.44	25.46	95.37	1.97617
2	11.41	42.33	125.7	19.98893
3	8.53	37.59	109.7	27.67409
4	9.92	47.13	114.57	25.29973
5	13.8	36.61	77.71	23.8952
6	11.82	40.25	93.28	18.28288

Table 4. Suggested optimized values for desirable higher sensitivity of the gas sensor system.

Number	Fe$_3$O$_4$ Additive	NO$_2$	Response Time	Sensitivity
1	16.06	36.48	38.24	99.99985
2	2.53	15.7	163.88	99.99991
3	4.07	24.2	57.59	99.99984
4	1.99	9.02	215.99	99.99991
5	8.36	22.53	49.91	99.99987
6	5.29	20.51	89.78	100.0001

Likewise, the first suggested optimum values of the Fe_3O_4 agent, NO_2 gas concentration, and response time in Table 4 are 16.06%, 36.48 mg. L^{-1}, and 38.24 s, respectively, which leads to an estimated sensitivity of 99.999%. The desirability of these suggested optimum values at a constant response time of 38.24 s is depicted in Figure 5b. The noise on the curve exposes that the system is remarkably sensitive to the gas concentration and Fe_3O_4 amounts. Although the optimum values of the estimated model contain high desirability, the significant sensitivity to the independent variables and low predicted R-squared value of the model increase the demand for employing AI techniques to better study the behavior of the system.
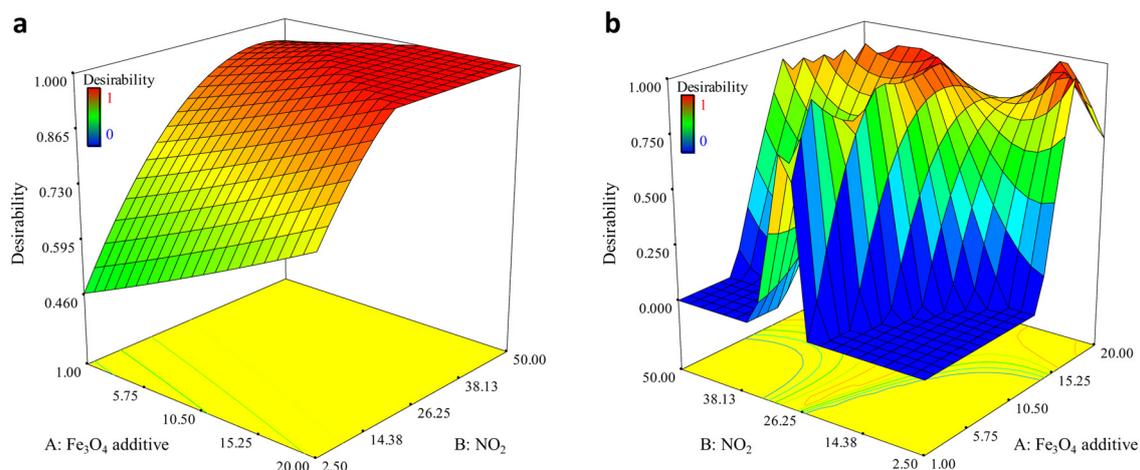
Figure 5. The desirability of the regression model for the prediction of (**a**) response time and (**b**) sensitivity.

3.2. Soft-Computing

According to RSM modelling, it can be found that the predicted-R^2 index in quadratic and cubic equations is less than 0.8 and, therefore, the importance of machine learning algorithm utilization is clearer than in the past.

The statistical outputs of machine learning calculations for the prediction of response time as one of the main targets of gas sensor performance assessment are summarized in Table 5. According to this table, it is clear that the best condition for the estimation of response time is related to M5Rules and Lazy.KStar algorithms with a 0.97 correlation coefficient. Furthermore, considering the two selected algorithms, the mean absolute error (MAE) of the M5Rules is less than that of the other, at 32.4. Therefore, the best performance is linked to the M5Rules algorithm. The equations of meta.RegressionByDiscretization, M5Rules, GP are demonstrated in Equations (S1)–(S3), respectively. However, in this research, for the evaluation of the highest efficiency of each algorithm, the tuning process of each calculation is performed based on the change in the K-fold value between the share of training and testing processes. According to Figure 6a, it can be understood that Lazy.KStar has appropriate precision for all the different folds, while the best performance of the GP algorithm appeared in more than 2-fold. Additionally, the behavior of both meta.RegressionByDiscretization and M5Rules is similar to that of GP.

Table 5. The statistical indicators of machine learning computations as per response time.

Response Time—Statistical Indicators	GP	Meta.RegressionByDiscretization	M5Rules	Lazy.KStar
Correlation coefficient	0.8306	0.9351	0.974	0.97
Mean absolute error	113.7162	49.2271	32.4638	37.5175
Root mean squared error	142.3465	63.0765	41.4937	47.3799
Relative absolute error	74.72%	32.34%	20.80%	24.04%
Root relative squared error	76.51%	33.90%	22.10%	25.23%

Similarly to the first function (response time), the second one (sensitivity), Lazy.KStar algorithm, with a correlation coefficient of 0.9888 (Table 6), has the best efficiency for the prediction of sensitivity. The summarized formulations of the machine learning computations for the estimation of sensitivity are demonstrated in Equations (S4)–(S6). According to Figure 6b, it is clear that for GP, M5Rules, and Lazy.KStar, the best condition occurs in 8-fold, while for meta.RegressionByDiscretization, the optimum efficiency is illustrated in 6-fold.

Table 6. The statistical indicators of artificial intelligence practices based on the sensitivity function.

Sensitivity—Statistical Indicators	GP	Meta.RegressionByDiscretization	Rules.M5Rules	Lazy.KStar
Correlation coefficient	0.7555	0.9299	0.9621	0.9888
Mean absolute error	21.06	10.6165	9.125	4.8856
Root mean squared error	27.64	18.0346	11.299	10.3669
Relative absolute error	60.19%	29.73%	26.32%	13.96%
Root relative squared error	61.81%	39.70%	25.48%	23.18%
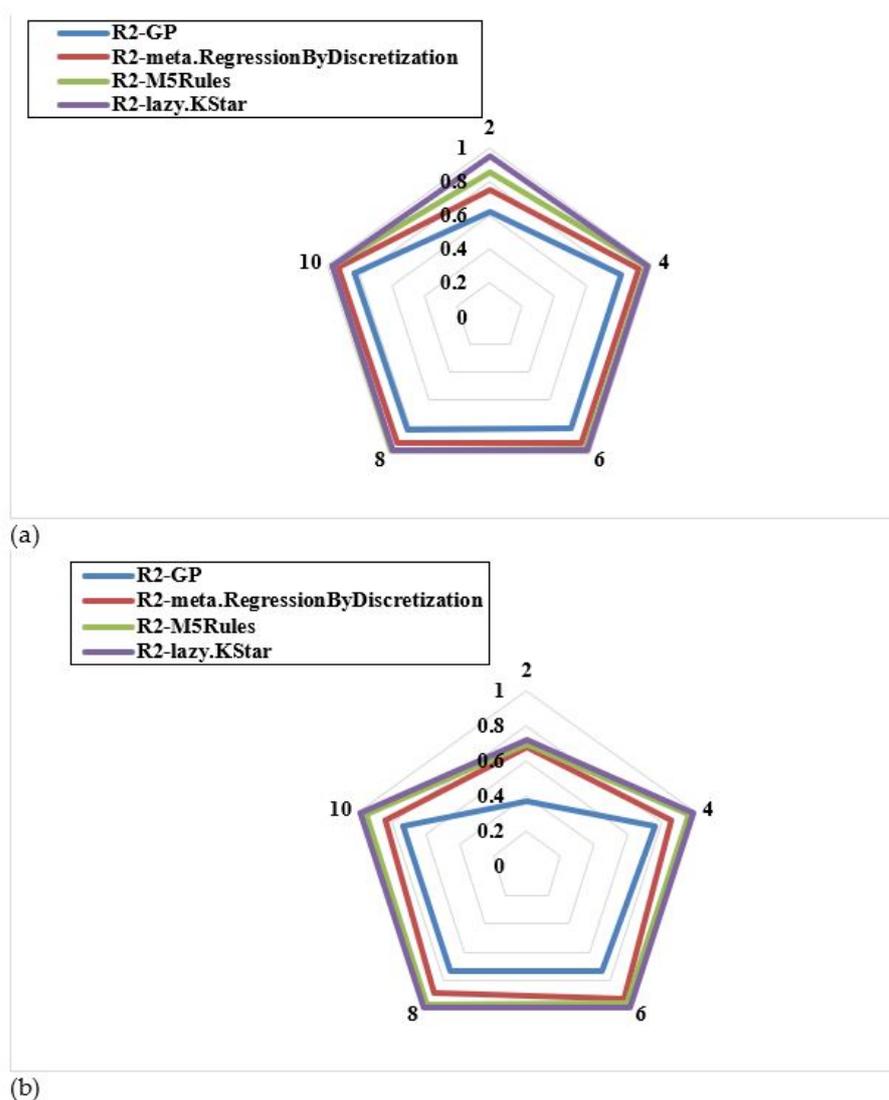
Figure 6. The R-squared fluctuations in different K-fold training numbers as per (**a**) response time and (**b**) sensitivity.

3.3. System Control of Gas-Sensor Operation by Jacobson

The conceptual model of a control system based on the Jacobson technique is demonstrated in Figure 7. In accordance with Figure 6, it can be understood that in the first step, after designing and synthesizing the gas-sensor in different conditions, the technology should be operated in the real field. It is clear that in the real execution of the sensor, the concentration of NO_2 gas varies as an air pollution compound.
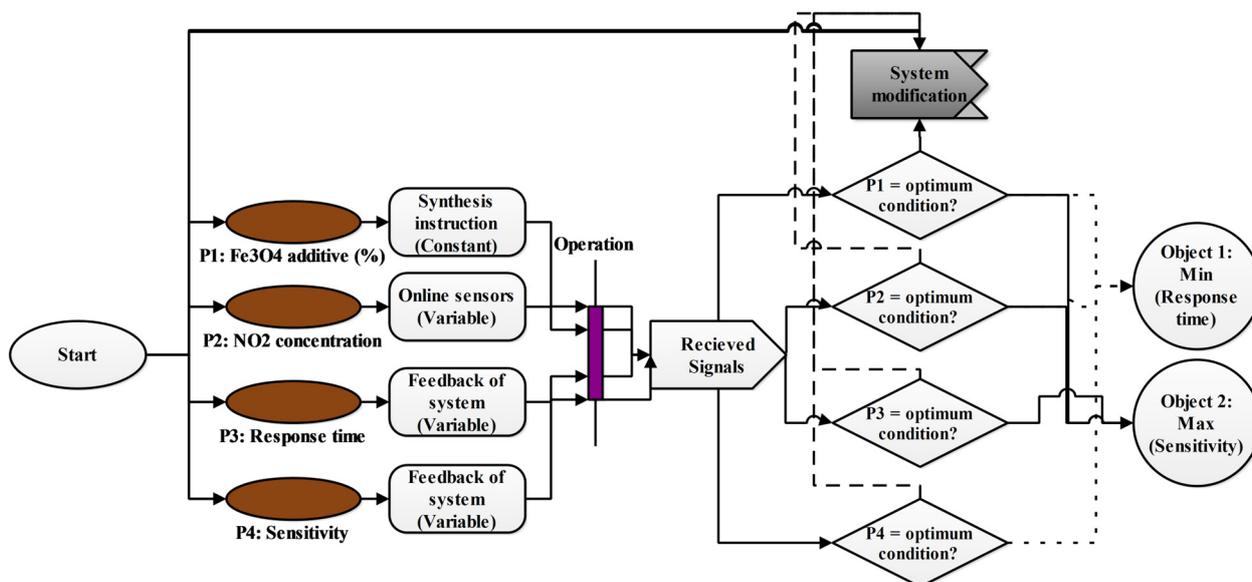
Figure 7. The conceptual model of the gas-sensor control system based on the Jacobson model in the present study.

Based on the operation process, sensitivity and response time are examined, and then, all signals (including P_1–P_4 features) can be evaluated under optimum conditions which are obtained from RSM analysis.

In the following, for object 1 (minimizing response time):

If P_1, P_2, and P_4 are set at optimum conditions, the goal is met.

Additionally, regarding the second objective (maximizing sensitivity):

If P_1–P_3 are equal to the optimal amounts, the purpose is satisfied.

Therefore, it can be concluded that with the application of the Jacobson technique and consideration of the outputs of RSM analysis, a logical control system can be implemented.

4. Conclusions

While gas sensing technology involves highly complicated design, invention, and operational features, artificial intelligence may aid in addressing these challenges. In this study, a soft-sensing method was introduced for the precise estimates of a NO_2 gas sensor using AI computations. However, before the prediction system, the process is optimized with the application of RSM, and the best estimation of Fe_3O_4 additive (%), input NO_2 (ppm), and sensitivity/response time is introduced according to forecasting both sensitivity and response time. The outputs of soft-computing demonstrated that the RSM has high efficiency for the sensitive analysis of effective factors and, also, the optimization of the NO_2 gas-sensing process based on both sensitivity and response time could be applied (with more than 0.98 R^2). Although the predicted R^2 of RSM as per the effective features is low (less than 0.8), this challenge can be addressed through artificial intelligence (more than 0.96). Finally, the conceptual modelling proved that the Jacobson technique has a high efficiency for the implementation of the gas-sensor control system. For future studies, the present research suggests the application of metaheuristics algorithms for the online optimization of the gas sensing process, whereas the deep learning algorithm application for the prediction of sensitivity/response time can be attractive for other investigations.

References

1. Ouyang, H.; Tang, X.; Kumar, R.; Zhang, R.; Brasseur, G.; Churchill, B.; Alam, M.; Kan, H.; Liao, H.; Zhu, T.; et al. Toward Better and Healthier Air Quality: Implementation of WHO 2021 Global Air Quality Guidelines in Asia. *Bull. Am. Meteorol. Soc.* **2022**, *103*, E1696–E1703. [CrossRef]
2. Hu, F.; Guo, Y. Health impacts of air pollution in China. *Front. Environ. Sci. Eng.* **2020**, *15*, 74. [CrossRef]
3. Zhang, Y.; Shi, T.; Wang, A.-J.; Huang, Q. Air Pollution, Health Shocks and Labor Mobility. *Int. J. Environ. Res. Public Health* **2022**, *19*, 1382. [CrossRef] [PubMed]
4. Chen, X.-Y.; Wang, X.-Z.; Liu, F.-J.; Zhang, G.-S.; Song, X.-J.; Tian, J.; Cui, H.-Z. Fabrication of porous Zn2TiO4–ZnO microtubes and analysis of their acetone gas sensing properties. *Rare Met.* **2020**, *40*, 1528–1535. [CrossRef]
5. Lin, Y.-C.; Cai, Y.; Huang, H.-Y.; Liang, D.; Li, J.; Tang, Y.; Hong, H.-C.; Yan, Q.; Huang, H.-D.; Li, Z. Air pollution and blood pressure in the elderly: Evidence from a panel study in Nanjing, China. *Heliyon* **2022**, *8*, e10539. [CrossRef]
6. Baravati, Z.A.; Khanjani, N.; Malakootian, M. Air Pollution and Mortality in the Elderly in Kerman, Iran. *Health Scope* **2021**, *10*, e105567. [CrossRef]
7. Chen, Y.; Cao, F.; Xiao, J.-P.; Fang, X.-Y.; Wang, X.-R.; Ding, L.-H.; Wang, D.-G.; Pan, H.-F. Emerging role of air pollution in chronic kidney disease. *Environ. Sci. Pollut. Res.* **2021**, *28*, 52610–52624. [CrossRef]
8. Sillmann, J.; Aunan, K.; Emberson, L.; Bueker, P.; Van Oort, B.; O'Neill, C.; Otero, N.; Pandey, D.; Brisebois, A. Combined impacts of climate and air pollution on human health and agricultural productivity. *Environ. Res. Lett.* **2021**, *16*, 093004. [CrossRef]
9. Chen, H.; Zeng, W.; Li, J.; Ma, T.; Liu, S.; Lei, G.; Gaiser, T.; Srivastava, A.K. Impact of Air Pollution on Maize and Wheat Production. *Ecol. Chem. Eng. S* **2022**, *29*, 237–256. [CrossRef]
10. Rovella, N.; Aly, N.; Comite, V.; Randazzo, L.; Fermo, P.; Barca, D.; de Buergo, M.A.; La Russa, M.F. The environmental impact of air pollution on the built heritage of historic Cairo (Egypt). *Sci. Total. Environ.* **2020**, *764*, 142905. [CrossRef] [PubMed]
11. Šetka, M.; Claros, M.; Chmela, O.; Vallejos, S. Photoactivated materials and sensors for NO_2 monitoring. *J. Mater. Chem. C* **2021**, *9*, 16804–16827. [CrossRef]
12. Laref, R.; Losson, E.; Sava, A.; Siadat, M. Empiric Unsupervised Drifts Correction Method of Electrochemical Sensors for in Field Nitrogen Dioxide Monitoring. *Sensors* **2021**, *21*, 3581. [CrossRef] [PubMed]
13. Patel, S.; Park, H.; Bonato, P.; Chan, L.; Rodgers, M. A review of wearable sensors and systems with application in rehabilitation. *J. Neuroeng. Rehabil.* **2012**, *9*, 21. [CrossRef]
14. Zhang, C.; Li, Y.; Liu, G.; Liu, K.; Wu, K. Room temperature NO_2 sensing properties of $ZnO_{1-\alpha}$ coating prepared by hydrogen reduction method. *Ceram. Int.* **2021**, *47*, 29873–29880. [CrossRef]

15. Krishna, K.G.; Parne, S.; Pothukanuri, N.; Kathirvelu, V.; Gandi, S.; Joshi, D. Nanostructured metal oxide semiconductor-based gas sensors: A comprehensive review. *Sens. Actuators A Phys.* **2022**, *341*, 113578. [CrossRef]

16. European Union. Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. *Off. J. Eur. Union* **2008**, *152*, 1–44.

17. European Parliament. *Sampling Points for Air Quality*; EU Publications: Luxembourg, 2019.

18. Xia, T.; Catalan, J.; Hu, C.; Batterman, S. Development of a mobile platform for monitoring gaseous, particulate, and greenhouse gas (GHG) pollutants. *Environ. Monit. Assess.* **2020**, *193*, 7. [CrossRef] [PubMed]

19. Chen, S.; Yuval; Broday, D.M. Re-framing the Gaussian dispersion model as a nonlinear regression scheme for retrospective air quality assessment at a high spatial and temporal resolution. *Environ. Model. Softw.* **2019**, *125*, 104620. [CrossRef]

20. Samad, A.; Mimiaga, F.E.M.; Laquai, B.; Vogt, U. Investigating a Low-Cost Dryer Designed for Low-Cost PM Sensors Measuring Ambient Air Quality. *Sensors* **2021**, *21*, 804. [CrossRef]

21. Russell, H.S.; Frederickson, L.B.; Kwiatkowski, S.; Emygdio, A.P.M.; Kumar, P.; Schmidt, J.A.; Hertel, O.; Johnson, M.S. Enhanced Ambient Sensing Environment—A New Method for Calibrating Low-Cost Gas Sensors. *Sensors* **2022**, *22*, 7238. [CrossRef]

22. Sankar, V.; Balasubramaniam, K.; Ramaprabhu, S. Invasive Species Prosopis juliflora Derived Carbon Biomass/SnO_2 based Hazardous NO_2 Gas Sensor. In Proceedings of the 2021 IEEE Sensors, Sydney, Australia, 31 October–3 November 2021; pp. 1–4. [CrossRef]

23. Gholami, P.; Rashidi, A.; Abbasabadi, M.K.; Pourkhalil, M.; Jahangiri, M.; Izadi, N. Synthesis and characterization of ZnO-functionalized multiwall carbon nanotubes nanocomposite as NOx gas sensor. *Res. Chem. Intermed.* **2020**, *46*, 3911–3927. [CrossRef]

24. Pérez, R.; Ayala, C.; Park, J.-Y.; Choi, J.-W.; Warner, I. Coating-Based Quartz Crystal Microbalance Detection Methods of Environmentally Relevant Volatile Organic Compounds. *Chemosensors* **2021**, *9*, 153. [CrossRef]

25. Mondal, B.; Meetei, M.; Das, J.; Chaudhuri, C.R.; Saha, H. Quantitative recognition of flammable and toxic gases with artificial neural network using metal oxide gas sensors in embedded platform. *Eng. Sci. Technol. Int. J.* **2015**, *18*, 229–234. [CrossRef]

26. Minh, V.T.; Moezzi, R.; Owe, I. Fuel economy regression analyses for hybrid electric vehicle. *Eur. J. Electr. Eng.* **2018**, *20*, 363–377. [CrossRef]

27. Spinelle, L.; Gerboles, M.; Villani, M.G.; Aleixandre, M.; Bonavitacola, F. Field calibration of a cluster of low-cost commercially available sensors for air quality monitoring. Part B: NO, CO and CO_2. *Sens. Actuators B Chem.* **2017**, *238*, 706–715. [CrossRef]

28. Zimmerman, N.; Presto, A.A.; Kumar, S.P.N.; Gu, J.; Hauryliuk, A.; Robinson, E.S.; Robinson, A.L.; Subramanian, R. A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. *Atmospheric Meas. Tech.* **2018**, *11*, 291–313. [CrossRef]

29. Saab, S.; Fu, Y.; Ray, A.; Hauser, M. A Dynamically Stabilized Recurrent Neural Network. *Neural Process. Lett.* **2021**, *54*, 1195–1209. [CrossRef]

30. Kusuma, V.; Privadi, A.; Budi, A.L.S.; Putri, V.L.B. Photovoltaic Power Forecasting Using Recurrent Neural Network Based on Bayesian Regularization Algorithm. *IEEE Trans. Neural Netw.* **2021**, *21*, 109–114. [CrossRef]

31. Mirikitani, D.T.; Nikolaev, N. Recursive Bayesian Recurrent Neural Networks for Time-Series Modeling. *IEEE Trans. Neural Networks* **2009**, *21*, 262–274. [CrossRef] [PubMed]

32. Aharoni, Z.; Rattner, G.; Permuter, H. Brief Announcement: Gradual Learning of Deep Recurrent Neural Network. In Proceedings of the International Symposium on Cyber Security, Cryptology, and Machine Learning, Beer-Sheva, Israel, 21–22 June 2018; pp. 274–277. [CrossRef]

33. Fang, X.; Wang, Y.L.; Murphey, D.W.; MacNeille, P. Specific Humidity Forecasting using Recurrent Neural Network. In Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN), Beijing, China, 6–11 July 2014; pp. 955–960.

34. Allende, H.; Moraga, C.; Nanculef, R.; Salas, R. Ensembles Methods for Machine Learning. *Pattern Recognit. Mach. Vis.* **2010**, *6*, 247–261.

35. Barbez, A.; Khomh, F.; Guéhéneuc, Y.-G. A machine-learning based ensemble method for anti-patterns detection. *J. Syst. Softw.* **2019**, *161*, 110486. [CrossRef]

36. Li, T.; Wang, Y.; Yuan, Q. Remote Sensing Estimation of Regional NO_2 via Space-Time Neural Networks. *Remote Sens.* **2020**, *12*, 2514. [CrossRef]

37. Ghahremanloo, M.; Lops, Y.; Choi, Y.; Yeganeh, B. Deep Learning Estimation of Daily Ground-Level NO_2 Concentrations from Remote Sensing Data. *J. Geophys. Res. Atmos.* **2021**, *126*, e2021JD034925. [CrossRef]

38. Beloconi, A.; Vounatsou, P. Bayesian geostatistical modelling of high-resolution NO_2 exposure in Europe combining data from monitors, satellites and chemical transport models. *Environ. Int.* **2020**, *138*, 105578. [CrossRef] [PubMed]

39. Pankratova, N.D.; Nedashkovskaya, N.I. A decision support system for evaluation of decision alternatives on basis of a network criteria model. In Proceedings of the 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kyiv, Ukraine, 29 May–2 June 2017; pp. 830–835. [CrossRef]

40. Zhang, H.; Zhao, H.; Li, H.; Chen, Y.; Ai, J.; Wang, Q. A Decision Support System of Premium Power Supply Investment. In Proceedings of the 2020 5th Asia Conference on Power and Electrical Engineering (ACPEE 2020), Chengdu, China, 4–7 June 2020; pp. 934–939.

41. Khodabakhshi, S.; Abbasabadi, M.K.; Heydarian, S.; Shirazi, S.G.; Marahel, F. Fe$_3$O$_4$ Nanoparticles as Highly Efficient and Recyclable Catalyst for the Synthesis of 4-Hydroxy-3-[aryloyl(benzamido)methyl]coumarin under Solvent-Free Conditions. *Lett. Org. Chem.* **2015**, *12*, 465–470. [CrossRef]
42. Hummers, W.S., Jr.; Offeman, R.E. Preparation of Graphitic Oxide. *J. Am. Chem. Soc.* **1958**, *80*, 1339. [CrossRef]
43. Abbasabadi, M.K.; Rashidi, A.; Khodabakhshi, S. Benzenesulfonic acid-grafted graphene as a new and green nanoadsorbent in hydrogen sulfide removal. *J. Nat. Gas Sci. Eng.* **2016**, *28*, 87–94. [CrossRef]
44. Rashidi, A.; Tavakoli, Z.; Tarak, Y.; Khodabakhshi, S.; Abbasabadi, M.K. One-pot and Environmentally Friendly Synthesis of New Spiroindolones Using Functionalized Multiwall Carbon Nanotubes as Powerful Catalyst. *J. Chin. Chem. Soc.* **2016**, *63*, 399–403. [CrossRef]
45. Abbasabadi, M.K.; Zand, H.R.E.; Khodabakhshi, S.; Gholami, P.; Rashidi, A. Synthesis of new functionalized reduced graphene oxide quantum dot composite for high-performance NO$_2$ gas sensor. *Res. Chem. Intermed.* **2021**, *47*, 2279–2296. [CrossRef]
46. Amdoun, R.; Khelifi, L.; Khelifi-Slaoui, M.; Amroune, S.; Asch, M.; Assaf-Ducrocq, C.; Gontier, E. The desirability optimization methodology: A tool to predict two antagonist responses in biotechnological systems: Case of biomass growth and hyoscy-amine content in elicited Datura starmonium hairy roots. *Iran. J. Biotechnol.* **2018**, *16*, e1339. [CrossRef]
47. Eibe, F.; Hall, M.; Witten, I. *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*, 4th ed.; Morgan Kaufmann: Burlington, MA, USA, 2016.
48. Rajalakshmi, A.; Vinodhini, R.; Bibi, K.F. Data Discretization Technique Using WEKA Tool. *IJCSET* **2016**, *6*, 293–298.
49. Song, N.; Zhu, M.; Zhou, G.; Guo, L.; Mu, Y.; Gao, J.; Zhang, K. Design and optimization of halbach permanent magnet array with rectangle section and trapezoid section. *Int. J. Eng. Trans. B Appl.* **2021**, *34*, 184–191. [CrossRef]
50. Arab, M.; Akbarian, H.; Gheibi, M.; Akrami, M.; Fathollahi-Fard, A.M.; Hajiaghaei-Keshteli, M.; Tian, G. A soft-sensor for sustainable operation of coagulation and flocculation units. *Eng. Appl. Artif. Intell.* **2022**, *115*, 105315. [CrossRef]
51. Mahmood, Y.; Hussein, M.A. Intrusion Detection System Based on K-Star Classifier and Feature Set Reduction. *IOSR J. Comput. Eng.* **2013**, *15*, 107–112.
52. Taghavian, H.; Ranaei-Siadat, S.O.; Kalaee, M.R.; Mazinani, S.; Ranaei-Siadat, S.E.; Harati, J. Optimizing the activity of immobilized phytase on starch blended polyacrylamide nanofibers-nanomembranes by response surface methodology. *Fibers Polym.* **2015**, *16*, 1048–1056. [CrossRef]
53. Taghavian, H.; Ranaei-Siadat, S.O.; Kalaee, M.R.; Mazinani, S. Investigation of the effects of starch on the physical and biological properties of polyacrylamide (PAAm)/starch nanofibers. *Prog. Biomater.* **2017**, *6*, 85–96. [CrossRef] [PubMed]
54. Harati, J.; Ranaei Siadat, S.O.; Taghavian, H.; Kaboli, S.; Khorshidi, S. Improvement in biochemical characteristics of glycosylated phytase through immobilization on nanofibers. *Biocatal. Agric. Biotechnol.* **2017**, *12*, 96–103. [CrossRef]