



OPEN

Intelligent route to design efficient CO₂ reduction electrocatalysts using ANFIS optimized by GA and PSO

Majedeh Gheytnazadeh¹, Alireza Baghban^{2✉}, Sajjad Habibzadeh^{1✉}, Karam Jabbour³, Amin Esmaeili⁴, Amin Hamed Mashhadzadeh⁵ & Ahmad Mohaddespour³

Recently, electrochemical reduction of CO₂ into value-added fuels has been noticed as a promising process to decrease CO₂ emissions. The development of such technology is strongly depended upon tuning the surface properties of the applied electrocatalysts. Considering the high cost and time-consuming experimental investigations, computational methods, particularly machine learning algorithms, can be the appropriate approach for efficiently screening the metal alloys as the electrocatalysts. In doing so, to represent the surface properties of the electrocatalysts numerically, *d*-band theory-based electronic features and intrinsic properties obtained from density functional theory (DFT) calculations were used as descriptors. Accordingly, a dataset containing 258 data points was extracted from the DFT method to use in machine learning method. The primary purpose of this study is to establish a new model through machine learning methods; namely, adaptive neuro-fuzzy inference system (ANFIS) combined with particle swarm optimization (PSO) and genetic algorithm (GA) for the prediction of *CO (the key intermediate) adsorption energy as the efficiency metric. The developed ANFIS–PSO and ANFIS–GA showed excellent performance with RMSE of 0.0411 and 0.0383, respectively, the minimum errors reported so far in this field. Additionally, the sensitivity analysis showed that the center and the filling of the *d*-band are the most determining parameters for the electrocatalyst surface reactivity. The present study conveniently indicates the potential and value of machine learning in directing the experimental efforts in alloy system electrocatalysts for CO₂ reduction.

Transportation, power plants, and energy-intensive industries are some of the industrial activities that commonly emit greenhouse gases including CO₂, CH₄, and NO_x. Around 75% of greenhouse gas emissions are carbon dioxide emissions¹, which contribute to the 1.5 °C rise in global temperature that is considered to be a reasonably large amount. As a result, during the past 20 years, economic sustainability has gained international attention, drawing experts in the field of the environment, decision-makers, and international organizations from many nations. The United Nations Framework Convention on Climate Change was founded in 1992 as a result of this phenomena. The Kyoto Protocol and the 2015 Paris Agreement were later created in 1997 and 2015, respectively, to fight global warming by regulating greenhouse gas emissions².

Nowadays, one of the significant challenges in the energy sector that should be effectively addressed is the development of sustainable energy resources. This is because the current energy conversion technologies cannot meet the energy requirement³. The electrochemical reduction of carbon dioxide is a necessary process that uses electricity from renewables (solar, wind, etc.) to produce fuels or value-added chemicals from water and CO₂⁴. The *CO is the primary intermediate in the reduction of CO₂. Indeed, as explained in Hori et al. study⁵, it is the

¹Surface Reaction and Advanced Energy Materials Laboratory, Chemical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran. ²Chemical Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Mahshahr Campus, Mahshahr, Iran. ³College of Engineering and Technology, American University of the Middle East, Kuwait City, Kuwait. ⁴Department of Chemical Engineering, School of Engineering Technology and Industrial Trades, College of the North Atlantic - Qatar, Doha, Qatar. ⁵Mechanical and Aerospace Engineering, School of Engineering and Digital Sciences, Nazarbayev University, 010000 Nur-Sultan, Kazakhstan. ✉email: Alireza_baghban@alumni.ut.ac.ir; sajjad.habibzadeh@mail.mcgill.ca

only molecule having single carbon, which shows a product spectrum similar to CO₂ on a copper electrode. In this process, the determining component in providing an efficient conversion relies on the electrocatalyst^{6,7}. Several experiments and computations have been performed to find and prepare new materials as the electrocatalyst for the CO₂ electroreduction^{8–13}.

Therefore, understanding the surface chemical reactivity and the bond-breaking and -forming on the electrocatalyst's surface is required to describe such a surface phenomenon. A significant volume of concepts has been established for the adsorption of the simple molecules onto the surface of transition-metal electrocatalysts^{14–18}, among which the Sabatier principle is a main comprehensive concept¹⁹. It explains that the crucial reaction intermediate should have enough adsorption strength to increase the activity of the electrocatalyst; too strong or weak binding causes difficult product desorption and inadequate reactant activation, respectively. Thus, the reactivity as the function of binding energies can be described as a volcano-shaped plot²⁰. Nørskov and co-workers, through a series of pioneering studies, claimed that the chemisorption of the adsorbate is severely dependent on the surface electronic structure (*d*-band theory²¹).

According to the *d*-band theory, the strength of the bond is given by the construction of the bonding and antibonding states, between the transition metal *d* states and the adsorbate valence states, and antibonding states energy comparative to the Fermi level (filling). The antibonding states are upper the *d* states, making the energy of the *d* states center a well initial representor of the bond strength. The greater *d*-band center in energy results in higher adsorbate–metal antibonding states, less occupation of antibonding states, and stronger adsorption bonds²². Although due to the recent improvements in electronic structure approaches (primarily density functional theory (DFT)), the *d*-band theory could provide understandings of the surface activity for the transition metals along with a minor group of alloys^{3,23–29}, it could not explain the measured activity in some cases^{30–32}. This could be probably since the spread in energy states was not considered. The definition of parameters such as the upper edge of the *d*-band (E_u) and the *d*-band width (W_d) by using Hilbert transform of the projected density of states (DOS) has improved the correlation between the activity and the *d*-band center³¹.

Inspired by the *d*-band chemisorption theory, the thought is to develop a computational relation between the electronic and intrinsic properties (descriptors) of clean surfaces and *CO adsorption energy. This helps to perform the material screening employing these descriptors, suppressing the costly and time-consuming experiments. While many linear relations between the *d*-band theory characteristics and the binding energy of the adsorbate have been published^{31–33}, machine learning (ML) techniques were used in the present study in order to incorporate the possible nonlinear relation between the parameters, raise the accuracy of the prediction, and be more comprehensive.

ML techniques have been widely used to explore and investigate diverse phenomena in chemical engineering^{34–41}. In the present study, Ma et al.⁴² predicted the *CO binding energy on various alloy systems using artificial neural networks (ANN) algorithm with root mean square error (RMSE) of 0.13 eV. In another study, Li et al.⁴³ obtained almost the same accuracy (RMSE ~ 0.12 eV) when using the same ML algorithm but geometric descriptors as inputs. Recently, Noh et al.⁴⁴ proposed an ML model with RMSE of 0.05 eV. Unlike the other studies, they used two non-ab Initio input properties to forecast the binding energy of *CO in different systems. Since several aspects of an accurate model are still under question, it is necessary to establish a new model that enhances the prediction's accuracy and decreases the model's uncertainty.

A fuzzy method is one of the soft computing methods which has a significant capacity in modeling complicated and nonlinear systems. Adaptive neuro-fuzzy inference system (ANFIS), a hybrid of ANN and the fuzzy inference system (FIS), can anticipate the nonlinear systems significantly^{45–48}. The suggested model is based on the combination of genetic algorithm (GA) and particle swarm optimization (PSO) to ANFIS. To our knowledge, this investigation is the first attempt to apply ANFIS–GA and ANFIS–PSO in CO₂ electroreduction process. In this research, the *CO adsorption energy in the CO₂ electroreduction process was anticipated by the ANFIS–PSO and ANFIS–GA models. The PSO and GA were interconnected to ANFIS to balance the model's complication and their generality capacity. Some intrinsic features and DFT-calculated electronic features of clean alloy surfaces, explicitly, parameters of the *d*-band distribution were used as inputs of the model.

Methodology

ANFIS background. The ANFIS structure used in this study includes five layers and thirteen inputs, where the Takagi–Sugeno system was employed as FIS. The fuzzy rules are presented as follows^{46,49}:

Rule *i* :

If x_1 and x_2 are A_1 and B_1 , respectively, then $f_i = k_1^i x_1 + k_2^i x_2 + k_0^i$

where k_1 , k_2 , and k_0 represent the function parameters of output (f) and A and B indicate the membership functions for inputs (x_1 and x_2). The basic arrangement of ANFIS is a feedforward system that comprises five layers having several functions⁵⁰ (Fig. 1). The learning structure of ANFIS is well discussed by several studies^{51–53}.

The ANFIS with five layers is explained⁵⁴:

In Layer 1, fuzzification, the whole of the nodes are supposed to be adaptive inputs.

$$O_i^1 = \mu A_i(x_1) \quad i = 1, 2, \dots, n \quad (1)$$

$$O_i^1 = \mu B_{i-2}(x_2) \quad i = 3, 4, \dots, n \quad (2)$$

where n and both of $\mu A_i(x_1)$ and $\mu B_{i-2}(x_2)$ are the number of fuzzy sets per input variables and functions of Gaussian membership, respectively.

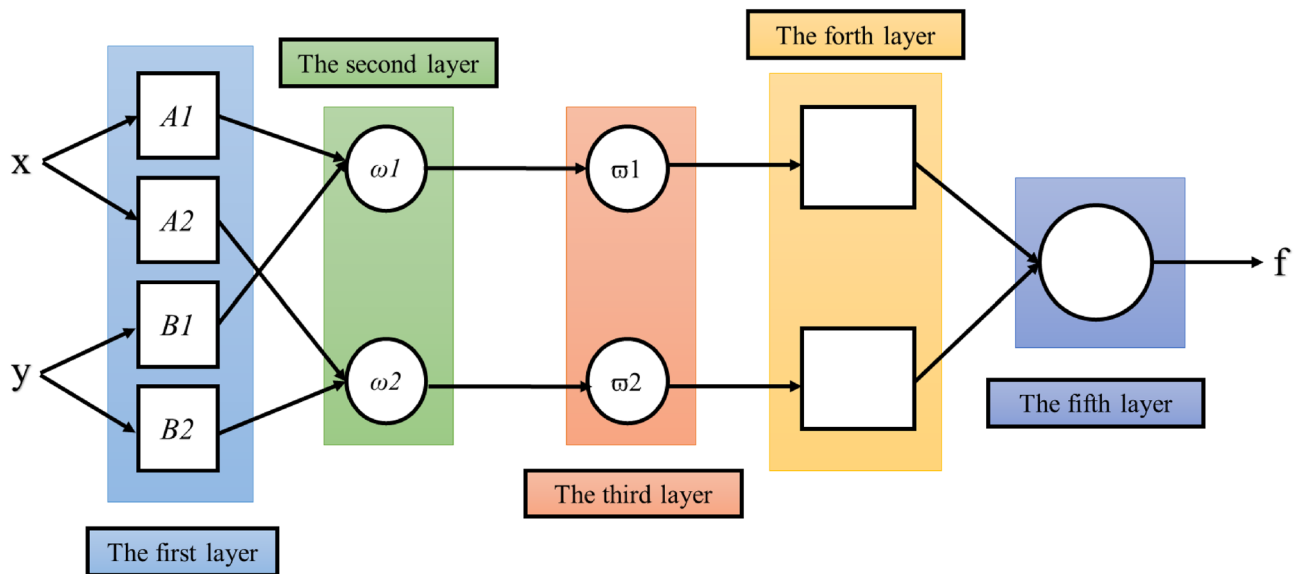


Figure 1. ANFIS structure.

Layer 2, which is the product layer, the output of this layer is the product of the input signal based on:

$$O_i^2 = \omega_i = \mu A_i(x_1)\mu B_i(x_2) \quad i = 1 \text{ or } 2 \tag{3}$$

Layer 3, named the normalized layer, a circle node is used to represent each node. The normalization function is defined as:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad i = 1 \text{ or } 2 \tag{4}$$

Layer 4, defuzzification, every node adapted with the following function:

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (k_1^i x + k_2^i y + k_0^i) \quad i = 1 \text{ or } 2 \tag{5}$$

where $(k_1^i x + k_2^i y + k_0^i)$ and \bar{w}_i are the variable sets of \bar{w}_i 's node and the output of Layer 3, respectively.

In Layer 5, which is the output layer, summation of the arrival signals creates the output model:

$$O_i^5 = \text{overall output} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad i = 1, 2 \tag{6}$$

Generally, an adaptive FIS comprises two diverse sections, the premise section and the consequent Section⁴⁷, optimized with various methods. A hybrid learning optimization method is developed, including the least squares method and the gradient descent method. In the present study, the fuzzy c-means (FCM) process is interconnected to ANFIS algorithm to forecast *CO adsorption energy.

FCM clustering. FCM is a method for data clustering where each point of data categorizes in clusters with different membership grades. FCM distributes a group of n vector $x_i, i = 1, 2, \dots, n$ into fuzzy categories and find a center for the cluster in each category such that the cost function is minimized according to the dissimilarity value.

By membership matrix U, the cost function can be calculated as:

$$J(U, v_1, \dots, v_K) = \sum_{i=1}^n \sum_{k=1}^K (u_{ik})^m d^2(x_i, v_K) \tag{7}$$

here $v_K = v_{ka}, k = 1, \dots, k, a = 1, \dots, p$ stands for the center of mass of kth cluster. In addition, m refers to the cluster's degree.

Algorithm should solve U and v_1, \dots, v_K by minimizing Eqs. (8 and 9).

$$u_{ik} = \frac{1}{\sum_{l=1}^K \left(\frac{d(x_i, v_k)}{d(x_i, v_l)} \right)^{\frac{2}{m}-1}} \tag{8}$$

$$v_k = \frac{\sum_{i=1}^n (U_{ik})^m x_i}{\sum_{i=1}^n (U_{ik})^m} \tag{9}$$

Minimizing task is carried out by the following cost function.

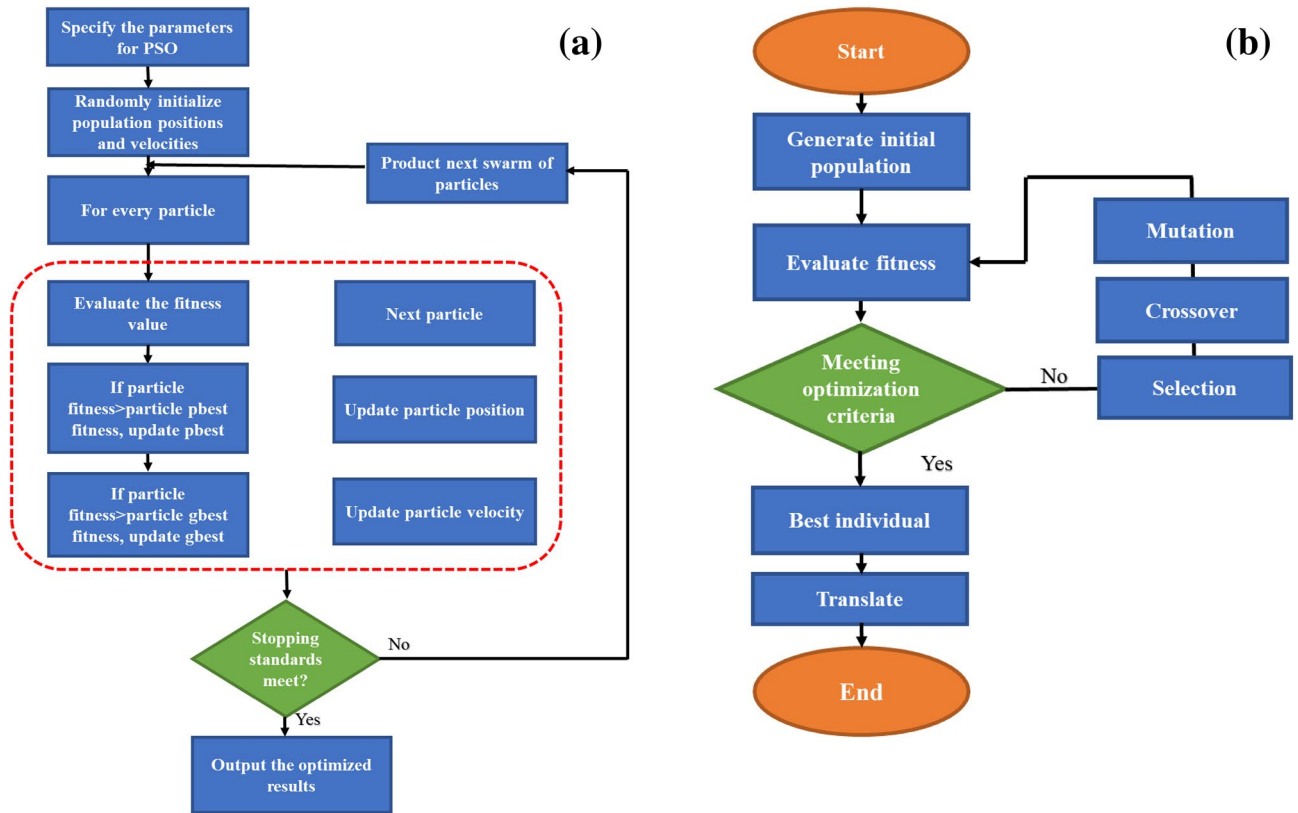


Figure 2. Scheme of the (a) PSO algorithm, (b) GA algorithm.

$$J(U) = \sum_{k=1}^K \left(\frac{\sum_{i=1}^n \sum_{j=1}^n ((u_{ik})^m (u_{jk})^m d_{ij})}{2 \sum_{s=1}^n (u_{sk})^m} \right) \tag{10}$$

Nikafshan Rad et al.⁴⁶ presented the FCM algorithm in detail.

PSO algorithm. PSO is a population-based optimization approach that randomly selects a population of particles or solutions, then seeks optima through updating generation in each iteration³⁹. Each particle is updated based on two particles: “pbest” and “gbest.” The “pbest” is the own best answer reached so far by particles, and the “gbest” is the total best solution obtained among the entire particles. Figure 2 shows the structure of PSO. In this scheme, random apportion velocities and positions are used at the first stage to start the initial population. The next step is to approximate each particle by regression analysis. When the stopping condition is satisfied by the best fitness rate of the particle, the process must be ended, and the factors should be reported. Suppose the fitness rate is not satisfactory for the ending criterion. In that case, the particles’ velocity and positions must be updated under two situations: In the first situation, if the particle fitness is larger than the gbest fitness, the related factors of gbest fitness should be updated. Second, if particle fitness is greater than the fitness of pbest, the pbest fitness factors should be updated. The other particles should be approximated through the second stage again.

Genetic algorithm. GA is a general accidental exploration tool for optimization based on the ideas of natural selection and genetics. It uses probabilistic transition rules instead of definite ones, which results in capacity to investigate big solution spaces. GA includes three stages: (a) population initialization, (b) operators of GA (selection, crossover, and mutation), and (c) evaluation^{49,55}.

- (1) *Population initialization:* In GA, each solution is named chromosome or string and is described via a series of different values. The strings (solutions) presenting each limitation or optimization issue necessities must be included in the initial population.
- (2) *Operators of GA:*
 - (1) *Selection:* It chooses solutions or individuals with high fitness values with a higher chance of surviving. A combination of Elitist and Tournament approaches is employed in this research through which the best solutions are chosen according to their fitness values and passed straight to the following generation.

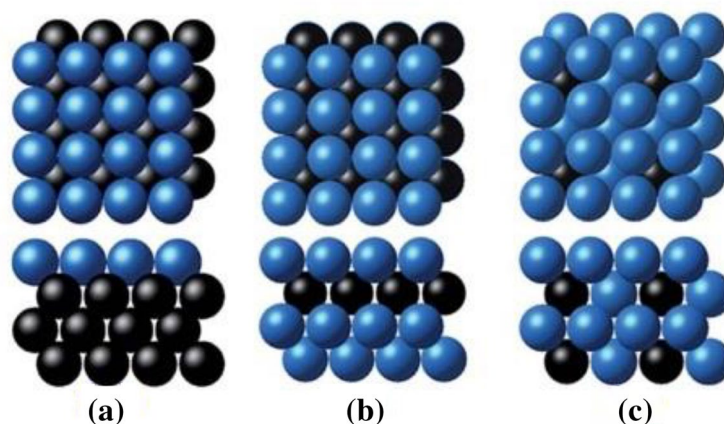


Figure 3. The {100}-terminated alloy surfaces in the form of (a) B@A, (b) AB@A, and (c) A₃B@A. The first rows indicate the top view, and the second rows show the side view of the structures⁴².

- (2) *Crossover*: In this one, two chosen individuals exchange part of their genes for making novel individuals for the following generation. Here, the scattered random method is used to create a new chromosome.
 - (3) *Mutation*: It makes a random change to the info within the strings or chromosomes. Occasionally, gene mutation occurs with a low possibility of transforming into new genes. The mutation operation increases the exploration ability of the search scheme to forbid to trap into local optima.
- (3) *Evaluation*: The function of fitness is applied to fit each individual solution.

The scheme of GA is illustrated in Fig. 2^{49,56}.

Modeling procedure

Data collection. In order to establish a relationship between the *CO binding energy and properties of electrocatalyst surfaces, a set of {100}-terminated bimetallic surfaces in the form of B@A, A-B@A, and A₃B@A have been considered as illustrated in Fig. 3. In the B@A structure, group VIII and IB metals (Cu, Au, Ni, Ag, Pt, and Pd) were chosen for A and B, but in the other structures, B contains the post-transition and the d-block metals. The required properties as input parameters were obtained through geometry optimization in DFT calculations. The input parameters are classified into two groups: primary features, electronic properties of the d-states distribution, and secondary properties, which are physical constants of the host metal. The secondary properties were used for a well description of the chemical bonding on a sequences of metal surfaces. The width (square root of the 2nd central moment, W_d), center (1st moment relative to the Fermi level, ϵ_d), filling (zeroth moment up to the Fermi level, f), the local Pauling electronegativity (χ), skewness (3rd standardized moment, γ_1), and kurtosis (4th standardized moment, γ_2) are primary features while work function (W), ionization potential (IE), square of adsorbate–metal interatomic d coupling matrix element (V_{ad}^2), spatial extent of metal d-orbitals (r_d), atomic radius (r_0), Pauling electronegativity (χ_0), and electron affinity (EA) are the secondary parameters. The 258 data points are taken from Ma et al. report⁴² and listed in Table S1. To develop the most precise model, randomly, 20% of the entire data was divided as the testing data to assay the model reliability and the remainder of them was used as the training data to study the *CO binding energy in CO₂ electroreduction systems.

Model evaluations. The statistical parameters (Eqs. 11–15) such as mean relative error (MRE), R², mean-square error (MSE), the standard deviation (STD), root-mean-square error (RMSE) were used to assess the model accuracy.

$$R^2 = 1 - \frac{\sum_{i=1}^n [x_i^{\text{predicted}} - x_i^{\text{experimental}}]^2}{\sum_{i=1}^n [x_i^{\text{predicted}} - x_m]^2} \quad (11)$$

$$STD = \sqrt{\frac{\sum_{i=1}^n (x_i^{\text{predicted}} - x_m)^2}{n}} \quad (12)$$

ANFIS-PSO		ANFIS-GA	
No. of inputs	258	No. of inputs	258
No. of output	1	No. of output	1
Membership function	Gaussian	Membership function	Gaussian
No. of cluster	5	No. of cluster	5
No. of tuning variables	140	No. of tuning variables	140
Population size	45	Population size	45
Iteration	1500	Iteration	1500
C1	1		
C2	2		

Table 1. Detail information of optimized ANFIS models.

$$MSE = \frac{1}{n} \sum_{i=1}^n \left(x_i^{predicted} - x_i^{experimental} \right)^2 \quad (13)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n \left(x_i^{predicted} - x_i^{experimental} \right)^2}{n}} \quad (14)$$

$$MRE = \frac{1}{n} \sum_{i=1}^n \frac{\left| x_i^{predicted} - x_i^{experimental} \right|}{x_i^{experimental}} \quad (15)$$

where n is the number of datapoints and m refers to mean value.

Model development procedure. As mentioned above, there are 13 input variables to predict target parameter. In this study, 5 clusters with Gaussian membership functions were used for construction of ANFIS model. Gaussian function comprises of 2 parameter that should be optimally determined during model development. Since there are 5 clusters and 13 inputs, there are 140 membership function parameters for optimization purpose. The optimum values of membership function parameters were determined by two evolutionary algorithms named PSO and GA. Setting of evolutionary algorithms was summarized in Table 1.

Accuracy of the collected data. Some suspected data or outliers in the data bank show inconsistent behavior with the rest of the data. They are probably generated because of the degree of the accuracy of the calculation and assumptions (or instrumental and human errors in the case of experimental data banks). It is vital to distinguish them because they can cause the wrong prediction for the developed model^{34,36}. To search the suspicious data and elevate the quality of data bank, the Leverage method is applied through which two parameters of critical leverage limit (H^*) and Hat matrix (H) are calculated as follows:

$$H = U \left(U^T U \right)^{-1} U^T \quad (16)$$

$$H^* = \frac{3j}{i + 1} \quad (17)$$

where i , j , and U are the number of the model parameters, the number of training data, and a matrix dimensional of $i \times j$, respectively. William's plot, the standardized residuals versus Hat values, is depicted in Fig. 4. In this plot, the reliable zone is defined as the region bounded between the standardized residuals of -3 to 3 and the critical leverage limit. According to Fig. 4, it can be clearly seen that most of the *CO binding energy values are placed in the valid area other than only 10 points among 258 data points, demonstrating the dataset is outstanding for training and testing models of ANFIS-PSO and ANFIS-GA.

Results and discussion

Analysis of sensitivity. Analysis of sensitivity is usually performed to study how the input parameters affect the output quantity³⁵. In this analysis, relevancy factor (r) indicates the most influential input parameter on *CO binding energy, which is calculated as:

$$r = \frac{\sum_{i=1}^n (X_{k,i} - \bar{X}_k) (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_{k,i} - \bar{X}_k)^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (18)$$

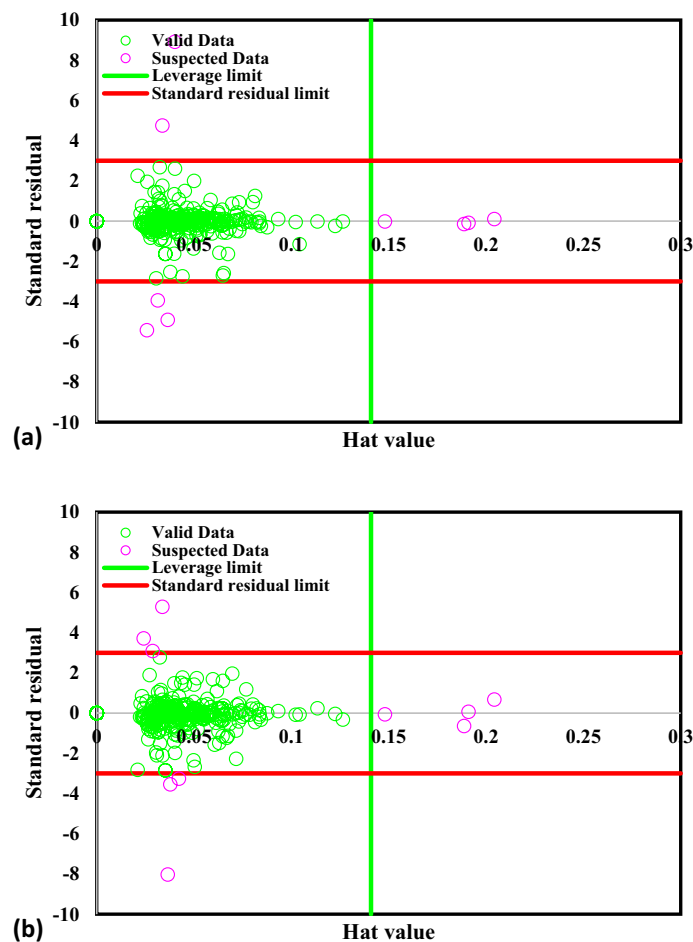


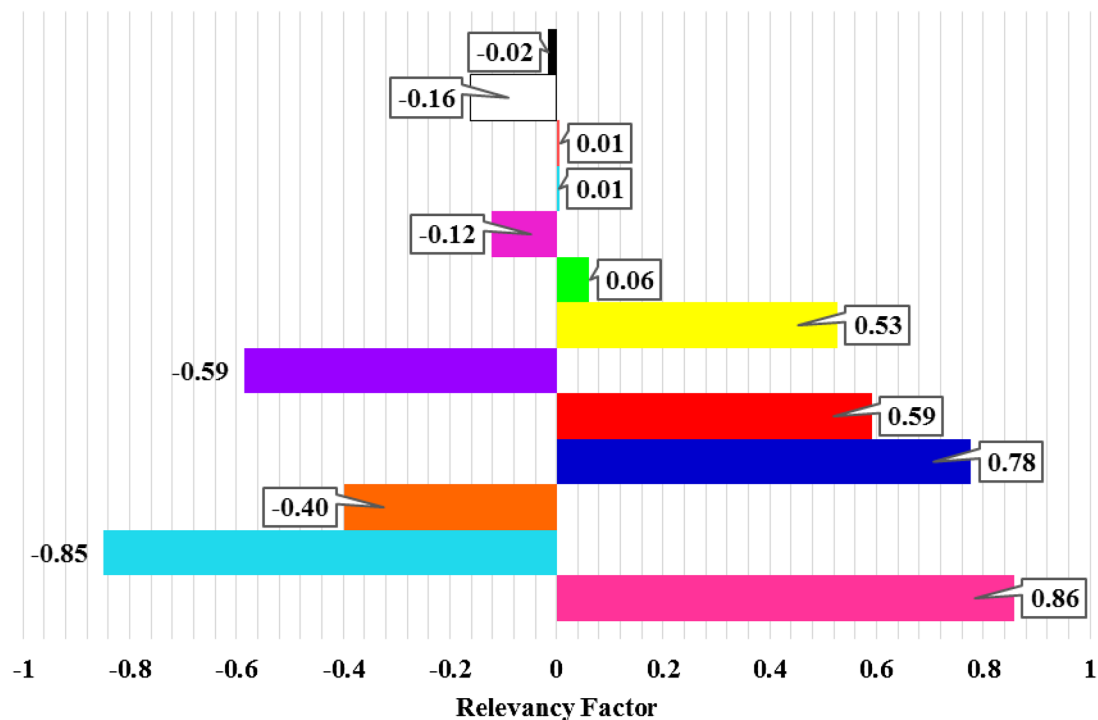
Figure 4. Detection of suspicious data for (a) PSO-ANFIS, (b) GA-ANFIS.

where $X_{k,i}$, \bar{X}_k , Y_i , \bar{Y} , and n indicate the 'k' th input parameter, the average of the input parameters, 'i' th output, the outputs average, and the number of all data points, respectively. Generally, r value changes between -1 to $+1$. The more absolute value of r shows the greater effect of the corresponding input on the output for each input. The negative and positive values denote that the high input the less and more in the output, respectively⁵⁷. According to Fig. 5, the *CO binding energy (consequently the reactivity of the metal surface) has direct relationship with the filling of the d -band, skewness of the d -band, Kurtosis of a d -band, atomic radius, spatial extent of d -orbitals, electron affinity, and Pauling electronegativity and has an inverse relationship with the center of the d -band, width of a d -band, work function, ionization potential, square of adsorbate-metal interatomic d coupling matrix element, and local Pauling electronegativity.

As expected from the d -band theory, the sensitivity analysis declared the filling and the center of the d -band are the most effective parameters by 0.86 and -0.85 relevancy factor, respectively. Also, the higher moments of the d -states distribution (such as skewness with r value of 0.78), which represent the d -band shape play a relatively important role in determining the reactivity of the metal surface. On the other hand, both of the electron affinity and the Pauling electronegativity have minimum effect on the *CO binding energy with r value of 0.01.

It is worth mention that parameters such as electron affinity and the Pauling electronegativity are intrinsic properties of active metal atoms which alter slightly across the periodic table but the most influencing parameters such as d -band center can be adjusted by ligand and strain engineering⁴².

Modeling results. To evaluate how precisely the suggested ANFIS-GA and ANFIS-PSO models are, the statistical factors are applied to determine the accordance between the data points and the anticipated values of the *CO binding energy, reported in Table 2. The ANFIS-GA and ANFIS-PSO models predicted the train data significantly excellent with R^2 of 0.994 and 0.995, respectively. The relative error quantities demonstrate the high precision of the proposed models in data training; in particular, the RMSE values of 0.0383 and 0.0411 for ANFIS-GA and ANFIS-PSO, respectively, are lower than the previous reports (0.13⁴², 0.12⁴³, and 0.05⁴⁴). This indicates the strength of the proposed models, which used the d -band theory electronic properties and the intrinsic features as inputs, the same data trained in⁴² with RMSE of 0.13. In addition to the forecast precision of the training data, the capability of the developed models to predict the unseen *CO binding energy data points is critically important. Therefore, both of the established models were investigated for the testing data. Again, both



■ Vad2 □ χ ■ χ_0 ■ EA (eV) ■ IE (eV) ■ rd (Å) ■ r0 (Å) ■ W (eV) ■ γ_2 ■ γ_1 ■ Wd (eV) ■ ϵ_d (eV) ■ f

Figure 5. Analysis of sensitivity of the input parameters for *CO adsorption energy on the alloy surfaces.

Model	Set	R ²	MRE (%)	MSE	RMSE	STD
ANFIS-PSO	Train	0.994	5.621	0.001688655	0.0411	0.0283
	Test	0.993	3.521	0.001807063	0.0425	0.0322
	Total	0.994	5.099	0.001718058	0.0425	0.0293
ANFIS-GA	Train	0.995	5.534	0.001467699	0.0383	0.0268
	Test	0.993	6.520	0.001770905	0.0421	0.0295
	Total	0.994	5.779	0.001542992	0.0421	0.0275

Table 2. The statistical parameters of proposed ANFIS-GA and ANFIS-PSO models.

of the ANFIS-GA and ANFIS-PSO models show close and great accuracy for prediction of the unseen *CO binding energy data, where MRE, R², MSE, STD, RMSE and are 3.521%/6.520%, 0.993/0.993, 0.00180/0.00177, 0.0322/0.0295, 0.0425/0.0421 and for ANFIS-PSO and ANFIS-GA algorithms, respectively.

For further confirmation the accuracy of the models, the data points and the anticipated values of the *CO binding energy are depicted by data indices in Fig. 6. As can be observed, the excellent agreement between the actual and estimated *CO binding energy amounts proves the excellent function of the established models. For both the ANFIS-GA and ANFIS-PSO models, the predicted values lines follow the actual data points precisely, which means these models have incredible capability to assess the reactivity of the metal alloy surfaces in terms of *CO adsorption energy.

The forecasted *CO binding energy values versus actual data values are depicted in Fig. 7. For both of the proposed ANFIS-GA and ANFIS-PSO models, the predicted data are precisely placed on their actual values so that their linear fitting shows correlation coefficients of more than 0.99. The 45° line is a criterion for the accuracy of the suggested models, where as shown in Fig. 7, the drawn fitting lines cross it precisely.

Moreover, the percentage of relative deviations between actual and estimated *CO binding energy data is shown in Fig. 8 for testing and training data sets of the developed ANFIS-GA and ANFIS-PSO models. The percentage mean relative deviations of the training and testing data obtained by the PSO-ANFIS model are 5.621% and 3.521%, respectively. Also, these values were 5.534% and 6.520% for training and testing datasets of ANFIS-GA model, respectively.

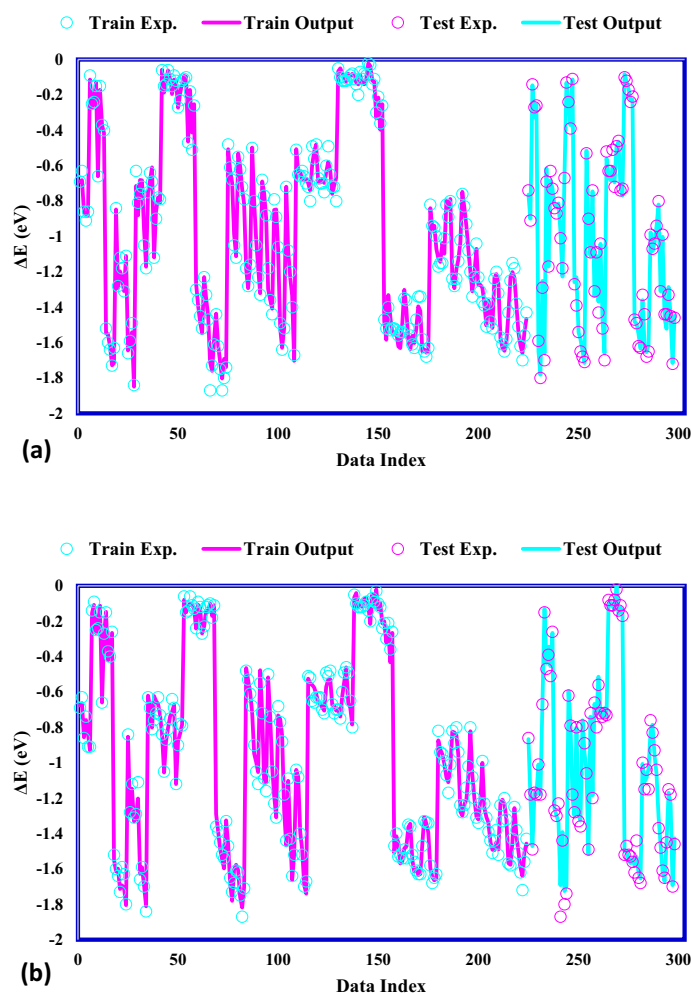


Figure 6. Comparison of actual and estimated values for (a) ANFIS-PSO, (b) ANFIS-GA.

Conclusions

The present study conveniently indicates the potential and value of machine learning in directing the experimental efforts in alloy system electrocatalysts for CO₂ reduction. Accordingly, two machine learning models of ANFIS-PSO and ANFIS-GA have been presented to anticipate the *CO binding energy on alloys electrocatalysts using electronic and intrinsic features. The required properties as input parameters were obtained through geometry optimization in DFT calculations.

The ANFIS-PSO and ANFIS-GA presented an excellent match between the *CO binding data and the predicted values with RMSE of 0.0411 and 0.0383, respectively, which is more precise than the other reported studies. In addition, it was found that the relative deviation of ANFIS-PSO model was 3.521%, while this value was 6.520% for the ANFIS-GA model, which indicates better ability of PSO to optimize ANFIS model. The percentage mean relative deviations of the training and testing data obtained by the PSO-ANFIS model are 5.621% and 3.521%, respectively. Also, these values were 5.534% and 6.520% for training and testing datasets of ANFIS-GA model, respectively. Moreover, the outlier detection technique was used to find outliers and it was shown by William's diagram. The sensitivity analysis confirmed the filling and the center of the *d*-band as the most influencing parameters for the electrocatalyst surface reactivity by 0.86 and -0.85 relevancy factor, respectively. This study's results and discussions can make it easier for scientists and chemical engineers to explore and examine new materials as electrocatalysts for the CO₂ electroreduction process.

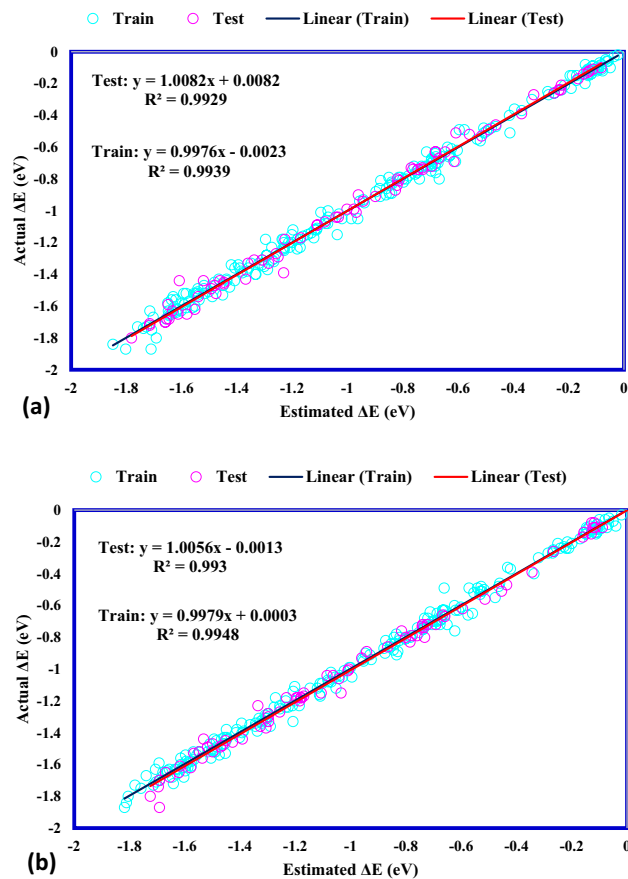


Figure 7. Cross plots for (a) ANFIS-PSO, (b) ANFIS-GA.

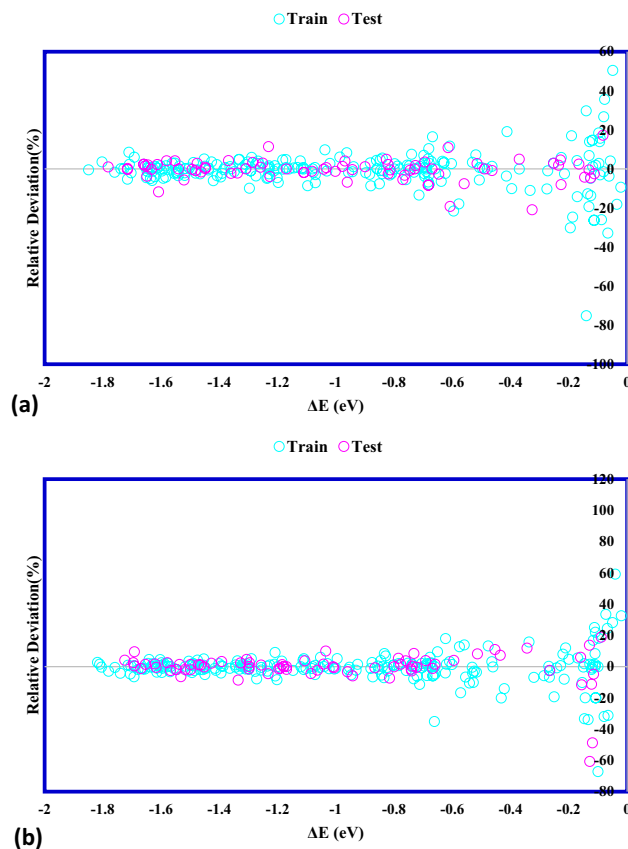


Figure 8. Comparison of experimental values and model outputs for (a) ANFIS-PSO, (b) ANFIS-GA.

Data availability

All data generated or analysed during this study are included in this published article [and its supplementary information files].

Received: 11 December 2021; Accepted: 30 November 2022

Published online: 02 December 2022

References

- Abbasi, F. & Riaz, K. CO₂ emissions and financial development in an emerging economy: An augmented VAR approach. *Energy Policy* **90**, 102–114 (2016).
- Kayani, G. M., Ashfaq, S. & Siddique, A. Assessment of financial development on environmental effect: Implications for sustainable development. *J. Clean. Prod.* **261**, 120984 (2020).
- Chen, A., Zhang, X., Chen, L., Yao, S. & Zhou, Z. A machine learning model on simple features for CO₂ reduction electrocatalysts. *J. Phys. Chem. C* **124**, 22471–22478 (2020).
- Laursen, A. B. *et al.* CO₂ electro-reduction on Cu₃P: Role of Cu(I) oxidation state and surface facet structure in C1-formate production and H₂ selectivity. *Electrochim. Acta* **391**, 138889 (2021).
- Hori, Y. Electrochemical CO₂ reduction on metal electrodes. *Mod. Asp. Electrochem.* https://doi.org/10.1007/978-0-387-49489-0_3 (2008).
- Maxwell, I. E. Driving forces for innovation in applied catalysis. *Stud. Surf. Sci. Catal.* **101A**, 1–9 (1996).
- Lewis, N. S. & Nocera, D. G. Powering the planet: Chemical challenges in solar energy utilization. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15729–15735 (2006).
- Baturina, O. A. *et al.* CO₂ electroreduction to hydrocarbons on carbon-supported Cu nanoparticles. *ACS Catal.* **4**, 3682–3695 (2014).
- Liu, X., Wang, Z., Tian, Y. & Zhao, J. Graphdiyne-supported single iron atom: A promising electrocatalyst for carbon dioxide electroreduction into methane and ethanol. *J. Phys. Chem. C* **124**, 3722–3730 (2020).
- Cao, L. *et al.* Mechanistic insights for low-overpotential electroreduction of CO₂ to CO on copper nanowires. *ACS Catal.* **7**, 8578–8587 (2017).
- Zhang, Q., Xu, W., Xu, J., Liu, Y. & Zhang, J. High performing and cost-effective metal/metal oxide/metal alloy catalysts/electrodes for low temperature CO₂ electroreduction. *Catal. Today* **318**, 15–22 (2018).
- Zhong, M. *et al.* Accelerated discovery of CO₂ electrocatalysts using active machine learning. *Nature* **581**, 178–183 (2020).
- Guo, Y. *et al.* Machine-learning-guided discovery and optimization of additives in preparing Cu catalysts for CO₂ reduction. *J. Am. Chem. Soc.* **143**, 5755–5762 (2021).
- Nilsson, A., Pettersson, L. G. M. & Nørskov, J. K. Chemical bonding at surfaces and interfaces. *Chem. Bond. Surf. Interfaces* <https://doi.org/10.1016/B978-0-444-52837-7.X5001-1> (2008).
- Nilsson, A. & Pettersson, L. G. M. Chemical bonding on surfaces probed by X-ray emission spectroscopy and density functional theory. *Surf. Sci. Rep.* **55**, 49–167 (2004).
- Nilsson, A. *et al.* The electronic structure effect in heterogeneous catalysis. *Catal. Lett.* **100**, 111–114 (2005).
- Van Santen, R. A. & Neurock, M. Molecular heterogeneous catalysis: A Conceptual and computational approach. *Mol. Heterog. Catal. Concept. Comput. Approach* <https://doi.org/10.1002/9783527610846> (2007).
- Ertl, G. Reactions at solid surfaces. *React. Solid Surf.* <https://doi.org/10.1002/9780470535295> (2010).
- Sabatier, P. Hydrogénations et déshydrogénations par catalyse. *Ber. Dtsch. Chem. Gesellschaft* **44**, 1984–2001 (1911).
- Medford, A. J. *et al.* From the Sabatier principle to a predictive theory of transition-metal heterogeneous catalysis. *J. Catal.* **328**, 36–42 (2015).
- Hammer, B. & Nørskov, J. K. Theoretical surface science and catalysis—calculations and concepts. *Adv. Catal.* **45**, 71–129 (2000).
- Nørskov, J. K., Abild-Pedersen, F., Studt, F. & Bligaard, T. Density functional theory in surface chemistry and catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 937–943 (2011).
- He, Y., Cubuk, E. D., Allendorf, M. D. & Reed, E. J. Metallic metal-organic frameworks predicted by the combination of machine learning methods and ab initio calculations. *J. Phys. Chem. Lett.* **9**, 4562–4569 (2018).
- Simón-Vidal, L. *et al.* Perturbation-theory and machine learning (PTML) model for high-throughput screening of parham reactions: Experimental and theoretical studies. *J. Chem. Inf. Model.* **58**, 1384–1396 (2018).
- Nørskov, J. K., Bligaard, T., Rossmeisl, J. & Christensen, C. H. Towards the computational design of solid catalysts. *Nat. Chem.* **1**, 37–46 (2009).
- Greeley, J. *et al.* Alloys of platinum and early transition metals as oxygen reduction electrocatalysts. *Nat. Chem.* **1**, 552–556 (2009).
- Back, S. & Jung, Y. Importance of ligand effects breaking the scaling relation for core-shell oxygen reduction catalysts. *Chem-CatChem* **9**, 3173–3179 (2017).
- Back, S., Kim, H. & Jung, Y. Selective heterogeneous CO₂ electroreduction to methanol. *ACS Catal.* **5**, 965–971 (2015).
- Stamenkovic, V. *et al.* Changing the activity of electrocatalysts for oxygen reduction by tuning the surface electronic structure. *Angew. Chem. - Int. Ed.* **45**, 2897–2901 (2006).
- Gajdoš, M., Eichler, A. & Hafner, J. CO adsorption on close-packed transition and noble metal surfaces: Trends from ab initio calculations. *J. Phys. Condens. Matter* **16**, 1141–1164 (2004).
- Xin, H., Vojvodic, A., Voss, J., Nørskov, J. K. & Abild-Pedersen, F. Effects of d-band shape on the surface reactivity of transition-metal alloys. *Phys. Rev. B Condens. Matter Mater. Phys.* **89**, 115114 (2014).
- Vojvodic, A., Nørskov, J. K. & Abild-Pedersen, F. Electronic structure effects in transition metal surface chemistry. *Top. Catal.* **57**, 25–32 (2014).
- Xin, H., Holewinski, A. & Linic, S. Predictive structure-reactivity models for rapid screening of pt-based multimetallic electrocatalysts for the oxygen reduction reaction. *ACS Catal.* **2**, 12–16 (2012).
- Gheytaanzadeh, M., Baghban, A., Habibzadeh, S., Mohaddespour, A. & Abida, O. Insights into the estimation of capacitance for carbon-based supercapacitors. *RSC Adv.* **11**, 5479–5486 (2021).
- Gheytaanzadeh, M. *et al.* Towards estimation of CO₂ adsorption on highly porous MOF-based adsorbents using gaussian process regression approach. *Sci. Rep.* **11**, 1–13 (2021).
- Ahmadi, M. H. *et al.* An insight into the prediction of TiO₂/water nanofluid viscosity through intelligence schemes. *J. Therm. Anal. Calorim.* **139**, 2381–2394 (2020).
- Baghban, A., Bahadori, M., Lemraski, A. S. & Bahadori, A. Prediction of solubility of ammonia in liquid electrolytes using least square support vector machines. *Ain Shams Eng. J.* **9**, 1303–1312 (2018).
- Baghban, A. & Khoshkham, A. Application of LSSVM strategy to estimate asphaltene precipitation during different production processes. *Pet. Sci. Technol.* **34**, 1855–1860 (2016).
- Bahadori, A. *et al.* Computational intelligent strategies to predict energy conservation benefits in excess air controlled gas-fired systems. *Appl. Therm. Eng.* **102**, 432–446 (2016).

40. Baghban, A., Abbasi, P. & Rostami, P. Modeling of viscosity for mixtures of Athabasca bitumen and heavy n-alkane with LSSVM algorithm. *Pet. Sci. Technol.* **34**, 1698–1704 (2016).
41. Baghban, A. Application of the ANFIS strategy to estimate vaporization enthalpies of petroleum fractions and pure hydrocarbons. *Pet. Sci. Technol.* **34**, 1359–1366 (2016).
42. Ma, X., Li, Z., Achenie, L. E. K. & Xin, H. Machine-learning-augmented chemisorption model for CO₂ electroreduction catalyst screening. *J. Phys. Chem. Lett.* **6**, 3528–3533 (2015).
43. Li, Z., Ma, X. & Xin, H. Feature engineering of machine-learning chemisorption models for catalyst design. *Catal. Today* **280**, 232–238 (2017).
44. Noh, J., Back, S., Kim, J. & Jung, Y. Active learning with non-: Ab initio input features toward efficient CO₂ reduction catalysts. *Chem. Sci.* **9**, 5152–5159 (2018).
45. Trivedi, R., Singh, T. N. & Gupta, N. Prediction of blast-induced flyrock in opencast mines using ANN and ANFIS. *Geotech. Geol. Eng.* **33**, 875–891 (2015).
46. Nikafshan Rad, H., Jalali, Z. & Jalalifar, H. Prediction of rock mass rating system based on continuous functions using Chaos-ANFIS model. *Int. J. Rock Mech. Min. Sci.* **73**, 1–9 (2015).
47. Hasanipanah, M., Amnieh, H. B., Arab, H. & Zamzam, M. S. Feasibility of PSO-ANFIS model to estimate rock fragmentation produced by mine blasting. *Neural Comput. Appl.* **30**, 1015–1024 (2018).
48. Alameer, Z., Elaziz, M. A., Ewees, A. A., Ye, H. & Jianhua, Z. Forecasting copper prices using hybrid adaptive neuro-fuzzy inference system and genetic algorithms. *Nat. Resour. Res.* **28**, 1385–1401 (2019).
49. Yang, H., Hasanipanah, M., Tahir, M. M. & Bui, D. T. Intelligent prediction of blasting-induced ground vibration using ANFIS optimized by GA and PSO. *Nat. Resour. Res.* **29**, 739–750 (2020).
50. Moayedi, H., Raftari, M., Sharifi, A., Jusoh, W. A. W. & Rashid, A. S. A. Optimization of ANFIS with GA and PSO estimating a ratio in driven piles. *Eng. Comput.* **36**, 227–238 (2020).
51. Jang, J. S. R. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Trans. Syst. Man Cybern.* **23**, 665–685 (1993).
52. Armaghani, D. J., Momeni, E., Abad, S. V. A. N. K. & Khandelwal, M. Feasibility of ANFIS model for prediction of ground vibrations resulting from quarry blasting. *Environ. Earth Sci.* **74**, 2845–2860 (2015).
53. Thomas, S., Pillai, G. N., Pal, K. & Jagtap, P. Prediction of ground motion parameters using randomized ANFIS (RANFIS). *Appl. Soft Comput. J.* **40**, 624–634 (2016).
54. Shahnazar, A. *et al.* A new developed approach for the prediction of ground vibration using a hybrid PSO-optimized ANFIS-based model. *Environ. Earth Sci.* **76**, 1–17 (2017).
55. Chen, X. & Wang, N. A DNA based genetic algorithm for parameter estimation in the hydrogenation reaction. *Chem. Eng. J.* **150**, 527–535 (2009).
56. Rezakazemi, M., Dashti, A., Asghari, M. & Shirazian, S. H₂-selective mixed matrix membranes modeling using ANFIS, PSO-ANFIS, GA-ANFIS. *Int. J. Hydrog. Energy* **42**, 15211–15225 (2017).
57. Baghban, A., Mohammadi, A. H. & Taleghani, M. S. Rigorous modeling of CO₂ equilibrium absorption in ionic liquids. *Int. J. Greenh. Gas Control* **58**, 19–41 (2017).

Author contributions

All authors collaborated in writing, data gathering, simulation, and conception.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-25512-8>.

Correspondence and requests for materials should be addressed to A.B. or S.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022