

Research Article

Barium Titanate Semiconductor Band Gap Characterization through Gravitationally Optimized Support Vector Regression and Extreme Learning Machine Computational Methods

Sunday O. Olatunji¹ and Taoreed O. Owolabi ²

¹College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

²Physics and Electronics Department, Adekunle Ajasin University, Akungba Akoko 342111, Ondo State, Nigeria

Correspondence should be addressed to Taoreed O. Owolabi; owolabitaoreedolakunle@gmail.com

Received 19 March 2021; Revised 17 April 2021; Accepted 20 April 2021; Published 27 April 2021

Academic Editor: Essam Houssein

Copyright © 2021 Sunday O. Olatunji and Taoreed O. Owolabi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Barium titanate (BaTiO_3) is a class of ceramic multifunctional materials with unique thermal stability, prominent piezoelectricity constant, excellent dielectric constant, environmental friendliness, and excellent photocatalytic activities. These features have rendered barium titanate indispensable in many areas of applications such as electromechanical devices, thermistors, multilayer capacitors, and electrooptical devices. The photocatalytic activity of barium titanate semiconductor is hindered by its large band gap and high rate of charge recombination. Doping of the parent barium titanate compound for band gap tuning is challenging and consumes appreciable time and other valuable resources. This present work relates the influence of foreign material incorporation into the parent barium titanate with the corresponding energy band gap by developing extreme learning machine- (ELM-) based models and hybridization of support vector regression (SVR) with gravitational search algorithm (GSA) using the structural lattice distortion that emanated from doping as model descriptors. The developed gravitationally optimized SVR (GSVR) is characterized with a low value of mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean square error (RMSE) of 0.036 eV, 1.145 eV, and 0.122 eV, respectively. The developed GSVR model outperforms ELM-Sine and ELM-Sig models using various performance evaluators. The developed GSVR model investigates the significance of iodine and samarium incorporation on the band gap of the parent barium titanate and the attained energy gaps conform excellently to the experimentally reported values. The demonstrated precision of the developed GSVR as measured from the closeness of its estimates with the measured values provides a quick and accurate method of energy gap characterization with circumvention of experimental stress and conservation of valuable time as well as other resources.

1. Introduction

Barium titanate represents a kind of ferroelectric ceramic multifunctional compound with promising potentials in lead-free devices such as thermistors, electromechanical devices, multilayer capacitors, and electrooptical devices [1, 2]. Thermal stability, prominent piezoelectricity constant, excellent dielectric constant, and environmental friendliness are unique features that facilitate the application of barium titanate as an electronic component as well as photocatalyst

[3]. The choice of barium titanate semiconductor becomes necessary since the most widely explored titanium-based semiconductors are of low quantum efficiency and resisted within UV illumination. In order to further enhance the photo response of barium titanate semiconductor under solar light irradiation, the construction of a heterojunction interface between two semiconductors has been identified as a unique method of enhancing charge separation and further optimizes band potentials in the compound [4, 5]. Non-metallic elemental incorporation into the parent crystal

structure has also been an effective method in the literature [6]. The purpose of this incorporation is to lower the band gap and surface charge alteration and further lead to a decrease in the charge-carriers recombination rate. Subsequently, after the doping, the incorporated dopants induce the following modifications on the barium titanate: (a) negative and positive shift of the conduction and valence band, respectively, which consequently narrows the band gap; (b) induction of high stability when doping with nonmetallic heavy atom; (c) dopant localized states being established within semiconductor band gap; and (d) intramolecular energy levels being perturbed which ultimately mixes pure triplet and singlet states and (f) distorts the crystal structure of the parent barium titanate semiconductor [6]. Extreme learning machine and combination of gravitational search algorithm (GSA) with support vector regression (SVR) computational methods are proposed in this work while the distortions as a result of the incorporated dopants serve as the predictors for band gap estimation.

The octahedron center in the structural lattice crystal of the host barium titanate holds Ti^{4+} ion while O^{2-} forms the composition of the center [3]. The oxygen octahedrons surround the void in which Ba^{2+} is situated. Doping, composite formation, or codoping of barium titanate strengthen the multifunctional capacities of the compound and promote its application in photocatalysis through band gap tailoring and reduction of charge recombination rate. There are two substitution sites associated with barium titanate which include Ba^{2+} (referred to as A-site) and Ti^{4+} site (called B-site) [5, 7]. Codoping and doping into these mentioned sites lead to improved ferromagnetic and optical properties. It worth mentioning that substitutions of foreign materials into these sites result in oxygen vacancies consequent upon oxidation state differences while the possible 3d electrons emanating from the dopants coupled with the oxygen vacancies strengthen the visible light absorption capacity of barium titanate. Additionally, the oxygen defects and the ions from the introduced dopants might reduce the electron-hole recombination rate through trap sites complementation which is one of the factors through which the photocatalytic activities could be enhanced [5]. The distortion due to the introduced dopants as contained in the lattice constants corresponding to the doped barium titanate are the inputs to the proposed hybrid GSVR and extreme learning-based models in this contribution.

Machine learning techniques have recently gained attention for solving complex and real-life problems that cannot be handled by conventional methods [8–12]. Support vector regression belongs to a class of these intelligent techniques which tackles nonlinear real-life problems using kernel function where the input data is transformed into feature space characterized with high dimensionality [13]. The transformation is carried out with the aid of kernel function which can be linear, sigmoid, polynomial, or Gaussian function. These unique characteristics have strengthened the applications of the SVR-based algorithms in many areas as solid-state physics [14, 15], laser spectroscopy [16, 17], and others [18, 19]. The parameters attributed to the SVR algorithm control the precision of the

model and efficient tuning of the parameters is the bedrock for achieving excellent performance. The gravitational search algorithm (GSA) with Newtonian physical principles is implemented for parameter tuning in this work.

An extreme learning machine is regarded as a feed-forward neural network with a single layer. The merits of this network as compared with the conventional neural network include high learning speed, excellent generalization performance, and computation of output weight during training procedures using the linear ordinary least square solution which is noniterative in nature [14, 20, 21]. The learning capacity of ELM extends beyond regression problems and can effectively handle clustering as well as classification problems. The universal approximation strength of the ELM algorithm allows approximation of a target function (which is continuous in nature) in the Euclidean space compact subset. These unique features of ELM-based models have rendered the algorithm indispensable in addressing many real-life problems [17, 22–24].

The structure of the remaining part of this work is arranged in the following order: Section 2 discusses the mathematical formulation and background of support vector regression, gravitational search algorithm, and the extreme learning machine. Section 3 presents the physical description of the employed dataset and its sources. The step-by-step computational strategies are highlighted in Section 3. Section 4 analyses the outcomes of the proposed models. Comparison of the estimates of the developed models with the experimentally determined values for characterizing the influence of different dopants on the band gap of barium titanate is also presented in Section 4. Section 5 discusses the conclusions of the research work.

2. Mathematical Formulations of the Implemented Algorithms

The formulations of the implemented support vector regression algorithm, as well as the extreme learning machine, are presented here. This section further presents the physical background of the implemented gravitational search optimization algorithm.

2.1. Background of the Support Vector Regression Algorithm. SVR aims at mapping crystal lattice parameters of doped BaTiO_3 to an elevated feature space of high dimensionality by the implementation of kernel mapping function $\beta(d)$ with nonlinearity features. The algorithm develops and constructs linear regression on all training samples $\{(d_1, E_1^b), (d_2, E_2^b), \dots, (d_m, E_m^b)\}$ where $d_j \in \mathfrak{R}^m$ and $E_j^b \in \mathfrak{R}^m$ in feature space after transformation [25–27]. The approximated regression function of the SVR algorithm is presented in

$$E^b(d, \omega) = (\omega \cdot \beta(d)) + b, \quad (1)$$

where $(\omega \cdot \beta(d))$ represents the dot product corresponding to the feature space while $\beta(d)$ stands for the mapping function.

The ε -insensitive loss function L_f presented in equation (2) is minimized in equation (3) purposely to ensure the minimization of Euclidian norm ($\|\omega^2\|/2$) and thereby results in a flat function:

$$L_f(E, E^b(d, \omega)) = \begin{cases} 0, & \text{if } |E - E^b(d, \omega)| \leq \varepsilon, \\ |E - E^b(d, \omega)| - \varepsilon, & \text{otherwise,} \end{cases} \quad (2)$$

$$R(E^b) = \frac{C}{m} \sum_{j=1}^m L_f(E, E^b(d, \omega)) + \frac{\|\omega^2\|}{2}. \quad (3)$$

$$\left\{ \frac{\|\omega^2\|}{2} + C \sum_{j=1}^m \{\xi_j + \xi_j^*\}, \text{ subjected to: } (E_j - \omega \cdot \beta(d) - b) \leq \varepsilon + \xi_j, (\omega \cdot \beta(d) + b - E_j) \leq \varepsilon + \xi_j^*, \xi_j, \xi_j^* \geq 0. \right. \quad (4)$$

The Lagrange multiplier coefficients (η_j and η_j^*) were obtained after solving the optimization problem presented in equation (4) to compute the following equation:

$$E^b(d, \eta_j, \eta_j^*) = \sum_{j=1}^{m_{SV}} (\eta_j - \eta_j^*) (\beta(d_j) \cdot \beta(d_i)) + b. \quad (5)$$

The obtained nonzero coefficients, as well as the crystal lattice parameters of doped BaTiO₃, make up the number of support vectors (m_{SV}) which are implementable on subsequent BaTiO₃ based compounds for energy band gap computation. The final SVR approximated relation is presented in the following equation:

$$E^b(d, \eta_j, \eta_j^*) = \sum_{j=1}^{m_{SV}} (\eta_j - \eta_j^*) (\beta(d \cdot d_j)) + b. \quad (6)$$

The expression for the kernel function that helps in data transformation is defined in the following equation:

$$\beta(d \cdot d_j) = \exp(-\lambda \|d_j - d\|^2), \quad (7)$$

where the kernel option is represented by λ .

The epsilon, kernel parameter, and the regularization factor are the hyperparameters of the algorithm that are required to be optimized for ensuring precise and robust SVR-based model. The optimization algorithm with the Newtonian working principle has been employed in this present work due to its robustness, interesting physical operational principle, quick convergence to optimality, and strong mathematical formulation.

2.2. Operational Physical Principle of the Gravitational Search Algorithm. The gravitational search algorithm is an optimization algorithm with a characteristic of being heuristic which navigates the probable solution search space through the principle of Newtonian gravitational force and combines the concept of mass and Newtonian

The parameter C in equation (3) is known as the regularization factor and trades off the flatness of the acquired function linking descriptors with target and the training error. This parameter is among the model hyperparameters that need to be well tuned to ensure a precise and robust model. Slack parameters (ξ_j, ξ_j^*) are incorporated into the convex optimization problem to limit the effect of some constraints that prevent the actualization of the flat model. The optimization expression is simplified as shown in the following equation [28, 29]:

gravitational law in general [30, 31]. A set of objects (called agents in this description), which are being affected by the gravitational pull and subsequently change their positions as well as the distances from other agents in accordance to their respective masses, are referred to as candidate solutions in the defined space. The agent's fitness is a function of its mass. Consider a search space containing m agent number, each with distinct masses and positions. Suppose that the agent j th in d -position is denoted by x_j^d . The position vector X_j of j th agent is presented in equation (8) after taking search space dimension N into consideration:

$$X_j = \{x_j^1, \dots, x_j^d, \dots, x_j^N\}, \quad \text{for } j = 1, 2, 3, \dots, m. \quad (8)$$

The gravitational force $F_{ji}^d(k)$ acting from the agent of mass i M_i^a (active gravitational mass) to that of mass j M_j^p (passive gravitational mass) at iteration time k in dimension d for a closed system of masses is presented in [14, 32] the following equation:

$$F_{ji}^d(k) = G(k) \frac{M_j^p(k) x M_i^a(k)}{R_{ji}(k) + \xi} (x_i^d(k) - x_j^d(k)), \quad (9)$$

where ξ represents a small value preventing zero denominators and promoting stochastic search to the algorithm.

The expression for the Euclidean distance R_{ji} between the two agents is presented in [33, 34] the following equation:

$$R_{ji}(k) = \|X_j(k), X_i(k)\|_2, \quad (10)$$

The gravitational constant $G(k)$ contained in equation (9) is expressed in the following equation:

$$G(k) = G_0 e^{\mu(k/T)}, \quad (11)$$

where μ and G_0 stand for the positive constant and initial value of gravitational constant, respectively. The optimum attainable iteration number is represented by T . The

expression for the total resultant force $F_j^d(k)$ acting on j th agent at iteration time k along d dimension is depicted by

$$F_j^d(k) = \sum_{i \neq j, i \in \text{Kbest}} \text{rand}_i^d F_{ji}^d(k), \quad (12)$$

where rand_i^d and Kbest , respectively, stand for random number within an interval range of $[0, 1]$ and the number of probable solutions generated with search space. At the commencement of the optimization procedures, Kbest equates the number of possible solutions and decreases gradually as the algorithm approaches the global solution. Equation (13) computes the acceleration $a_j^d(k)$ of j th agent along d dimension at time k :

$$a_j^d(k) = \frac{F_j^d(k)}{M_{jj}(k)}, \quad (13)$$

where M_{jj} represents the inertial mass.

The velocities of the agents with gravitational law description are updated in accordance with the following equation:

$$V_j^d(k+1) = \text{rand}_j x V_j^d(k) + a_j^d(k), \quad (14)$$

where $V_j^d(k)$ represents the velocity of j th agent along d dimension at time k .

Similarly, the position of the agent during the exploitation and exploration search of the defined search space is presented in equation (15) while the assumptions contained in equation (16) hold [35, 36]:

$$x_j^d(k+1) = x_j^d(k) + V_j^d(k+1), \quad (15)$$

$$M_j^a = M_j^p = M_{jj} = M_j, \quad j = 1, 2, \dots, m. \quad (16)$$

The fitness of the agent in the search space was calculated through the incorporation of equation (17) in equation (18):

$$Q_j(k) = \frac{\text{fit}_j(k) - \text{worst}(k)}{\text{best}(k) - \text{worst}(k)}, \quad (17)$$

$$M_j(k) = \frac{Q_j(k)}{\sum_{j=1}^m Q_j(k)}, \quad (18)$$

where $\text{worst}(k)$, $\text{fit}_j(k)$, and $\text{best}(k)$ represent the worst (highest value of root mean square error) fitness of j th agent (root mean square error of j th agent) and the best agent (lowest value of root mean square error), respectively. The mathematical expressions for the best and worst agent are, respectively, shown in the following equations:

$$\text{best}(k) = \min \text{fit}_i(k), \quad i \in \{1, \dots, m\}, \quad (19)$$

$$\text{worst}(k) = \max \text{fit}_i(k), \quad i \in \{1, \dots, m\}. \quad (20)$$

2.3. Extreme Learning Machine Mathematical Background. Typically, there are three distinct layers in the ELM operational principle which include the input layer, hidden layer, and output layer [20]. The input biases and the input layer

weights are randomly generated and cease to partake in the training process furthermore. Consider the crystal lattice parameters $L_p \in \mathbb{R}^{n \times d}$, $L_p = (L_1, \dots, L_n)^T$ as input data for modeling band gap of doped barium titanate while d and n are the feature and sample size, respectively. During training, the available experimental crystal lattice parameters of doped barium titanate are mapped to ELM random space with N -dimensionality. The output of the hidden layer is computed as presented in equation (21) after g nonlinear transformation [16, 24]:

$$h_j(L_p) = g(L_p^T \chi_j + b_j), \quad j \in [1, N]. \quad (21)$$

The activation function g might be sine, sigmoid, linear, polynomial, and triangular basis, among others. The output layer of the network is computed using the following equation:

$$E_g^{\text{predicted}}(L_p) = \sum_{j=1}^N \varphi_j h_j(L_p) = h(L_p)^T \phi = \tau, \quad (22)$$

where $h(L_p) = (h_1(L_p), \dots, h_N(L_p))^T$, $\phi = (\phi_1, \dots, \phi_N)^T$ = output weight, and τ = approximated energy band gap (E_g) of barium titanate.

In order to determine the coefficients ϕ of the output layer, the condition contained in the following equation needs to be satisfied:

$$\phi = \text{argmin}_{\phi} \|E_g^{\text{predicted}}(L_p) - E_g\|_2^2, \quad (23)$$

where $E_g = (E_1, \dots, E_m)^T$ is the energy band gap matrix of the input crystal lattice matrix L_p .

The ELM approximated function $E_g^{\text{predicted}}(L_p) = \tau$ represents the approximation of the measured band gap E_g matrix. Introduction of $\mathbf{H} \in \mathbb{R}^{m \times N}$ further simplifies the problem as presented in the following equation:

$$\mathbf{H} = \begin{pmatrix} h_1(L_{p1}) & \cdots & h_N(L_{p1}) \\ \cdots & \ddots & \cdots \\ h_1(L_{pm}) & \cdots & h_N(L_{pm}) \end{pmatrix}. \quad (24)$$

Therefore, equation (23) can be further minimized as presented in the following equation:

$$\phi = \text{argmin}_{\phi} \|\mathbf{H}\phi - E_g\|_2^2. \quad (25)$$

Solving equation (25) through generalized Moore–Penrose inverse completes the training processes of the ELM algorithm.

3. Computational Hybridization of the Algorithms and Data Acquisition

The computational strategies of the proposed hybrid model and those of ELM-based models are discussed in this section. The acquisition of the employed dataset as well as the physical connection between the predictors and the energy gap of doped barium titanate is also presented.

3.1. Data Physical Description and Acquisition. The structural lattice parameters of barium titanate semiconductors doped with different materials and their corresponding band gap employed for developing the proposed GSVR model are extracted from the literature [2–7, 37–40]. Replacement of Ti^{4+} ions in the parent barium titanate by the equivalent ions (say Sn^{4+}) from the dopants results in lower angle peak shifting with characteristic unit cell expansion due to ionic radii difference [41]. The expansion and contraction in the lattice parameters due to the octahedral coordination incorporation serves as the predictors to the developed model. The predictors and the energy gap of the doped barium titanate are analyzed statistically and the outcomes of the analysis are shown in Table 1. The inference about the entire content of the dataset can be deduced from the obtained values of the minimum, maximum, and mean. The information regarding the consistence in the band gap determination from measurement to measurement can be inferred from the presented standard deviation in the table.

The correlation coefficients are also presented to investigate the possibility of linear modeling techniques in handling the band gap prediction since the extent and degree of the linear relationship are controlled by the value of the correlation coefficient. From the obtained coefficients of correlation, it is observed that linear modeling methods would perform poorly in establishing a link between the energy gap semiconductors barium titanate and Structural parameters.

3.2. Computational Strategies of the Developed Hybrid Model.

Generation of support vectors from support vector regression algorithm was conducted within computing MATLAB environment while the hybridized gravitational search algorithm aims at optimizing the hyperparameters associated with SVR purposely to attain a precise and robust model. The entire dataset extracted from forty barium titanate based compounds were partitioned into training and testing set in the ratio of 8:2 where 80% of the dataset (thirty-two compounds) was employed for support vector acquisition and generation while the testing set takes the remaining 20% (equivalent to eight compounds). Before the described data partitioning, the dataset was randomized so that the data points can be evenly and uniformly distributed. The step becomes necessary for unbiased and efficient computation. The procedures and methodologies for computational hybridization of SVR with gravitational search algorithm are described in a stepwise manner as follows:

Step I. Definition of the search space and population initialization: the search space that houses the limits of the probable solutions to the hyperparameters (epsilon regularization factor and kernel option) is defined as [1000, 0.1, 0.1] and [1, 0.001, 0.001] for upper and lower limits, respectively. This means the upper and lower limits of the regularization factor were set at 1000 and 1, respectively, while those of epsilon were set at 0.1 and 0.001, respectively. The epsilon and kernel option span over the same search space. A definite number of populations were initiated within the defined search

TABLE 1: Statistical results of the utilized set of data.

Parameters	Band gap (ev)	a (Å)	b (Å)	c (Å)
Maximum	3.7600	4.0134	4.0134	4.0400
Minimum	2.3000	3.9400	3.9400	3.9895
Mean	3.0320	3.9828	3.9835	4.0165
Correlation coefficient		-10.85	-16.03	-19.88
Standard deviation	0.3671	0.0197	0.0201	0.0119

space while the significance of this number on the convergence of the model was further investigated.

Step II. Completion of support vector regression implementation circle for fitness evaluation: the fitness of each of the agents was evaluated by implementing support vector regression algorithm on the training set of data while the root mean square error between the experimental band gap and the predicted value serves as the measure of the fitness. The fitness evaluation is summarized as follows:

- Kernel function selection: a function was selected for data transformation among available pool of functions such as radial basis function (Gaussian function), linear, sigmoid, and polynomial.
- Incorporation of GSA agents and training dataset into the chosen transformation function: the contents of each agent of GSA were incorporated with the mapping function to train the SVR algorithm using the training set of data. The number of the initial population of the agent corresponds to the number of developed the SVR-based model while the root mean square error between the measured band gap and the predicted value determines the fitness of each of the agents (developed model).
- Fitness evaluation: the fitness of each of the agents was ranked from highest to the lowest with the agent that is characterized with the lowest root mean square error as the best fit while the agent with the highest root mean square error as the worst of the population. Equation (17) computes the best and the worst of the agents in the population. This evaluation was carried out using a testing set of data through the implementation of the acquired support vectors during the training phase of model development.

Step III. Update of GSA important parameters: the gravitational constant, worst agent of the population, and best agent are updated using equations (11), (19), and (20), respectively.

Step IV. Computation of inertial mass and the acceleration with which each of the agents is approaching convergence: from the results of the fitness obtained in equation (17), the inertial mass of each of the agents was computed using equation (18). The acceleration of each of the agents was calculated using equation (13).

Step V. Updating the position and velocity of the agent for population replacement: after the selection of the

best agent in the initial iteration, the entire population is replaced with the new population through velocity and position update using equations (14) and (15), respectively.

Step VI. Stopping conditions: the algorithm goes through a repeating circle from Step II to Step V until the number of iteration attains 100 or the root mean square error attains zero value.

3.3. Extreme Learning Machine Based Model Computational Strategy. The entire forty barium titanate based compounds doped with different materials were randomized and partitioned into training and testing set in the ratio of 4 : 1. The following steps were taken after the randomization and partitioning stages of model development.

Step I: the biases (b) and input weights (φ) were randomly generated using a pseudorandom number generator through seeding in MATLAB environment.

Step II: activation function was selected from a pool of functions such as sine function, sigmoid function, linear function, and triangular basis function. Hidden neuron was also selected as 1.

Step III: computation of hidden layer output matrix using equation (22) coupled with training set of data.

Step IV: determination of the coefficients of output layer through the implementation of equation (23).

Step V: computation of output layer weights through Moore–Penrose inverse implementation.

Step VI: for a given activation function, repeat Step II to Step V with the number of hidden neurons varying from two to one hundred. Compute the performance measuring parameters (root mean square error, mean absolute error, and correlation coefficient) for each of the models and save the parameters of the best model. The parameters of the best model include the input weights, biases, outputs weights, number of hidden neurons, and the activation function.

Step VII: repeat Step II to Step VII for each of the available activation functions.

Step VIII: evaluate each of the models using a testing set of data by computing the performance measuring parameters and save the model parameters.

The computational flowchart for the developed ELM-based model is presented in Figure 1.

4. Results and Discussion

The results of each of the developed models and their comparison are presented in this section. A comparison of the measured and estimated band gap is also presented.

4.1. Exploration and Exploitation of Search Space for Different Population Sizes. The significance of the exploitation and exploration capacities of the developed GSVR model on the number of agents within the search space is presented in

Figure 2. The exploration capacity of the algorithm becomes significant before locating the most probable region of the search space while exploitation capacity measures the probing efficiency of the algorithm around the global solution. At a different number of agents, the algorithm converges to the same global solution. This kind of convergence justifies the robustness of the developed hybrid model.

The optimum values of SVR hyperparameters as obtained from GSA are presented in Table 2. The table also presents the optimum hyperparameters of ELM-based models.

4.2. Performance Comparison between ELM-Based Models and GSVR Model. The developed hybrid GSVR and ELM-based models were evaluated during the training phase of model development and the results of the evaluation are presented in Figure 3. Figure 3(a) presents the evaluation on the basis of correlation coefficient (CC), Figure 3(b) compares using the mean absolute error (MAE) as a yardstick for performance evaluation while Figure 3(c) presents the comparison using root mean square error (RMSE) metrics.

The developed GSVR model outperforms ELM-Sine and ELM-Sig model with performance improvement of 68.34% and 68.33%, respectively, using CC as performance evaluator as presented in Figure 3(a) while performance enhancement of 99.96% and 99.97% were, respectively, obtained when the developed GSVR model is compared with ELM-Sine model on the basis of MAE and RMSE as presented in Figures 3(b) and 3(c). The performance superiority of the developed GSVR model over the ELM-based models can be attributed to the uniqueness of the hybridization and the strong structural risk minimization mathematical principle upon which the SVR algorithm is built. The comparison of the performance of the developed model during the testing phase of model development is presented in Figure 4.

The developed GSVR model performs better than ELM-Sine and ELM-Sig models on the basis of CC with performance improvement of 4.27% and 4.25%, respectively, as depicted in Figure 4(a). On the basis of RMSE as shown in Figure 4(b), the developed GSVR model shows the superior performance of 24.90% to the ELM-Sine model. The training phase of the GSVR model shows better performance of 7.92%, 99.96%, and 99.94% than the testing phase of model development using CC, MAE, and RMSE as performance evaluator, respectively. The testing phases of the ELM-based models outperform the training phase as can be observed in Figures 3 and 4. Performance improvements of 64.37%, 29.41%, and 44.55% were, respectively, obtained when comparing the performance of the testing phase of the ELM-Sine model with the training stage of model development on the basis of CC, MAE, and RMSE, respectively. The overall comparison for the entire dataset is presented in Figure 5.

Outstanding performances of 86.59% and 86.60% were, respectively, obtained while comparing the superiority of the GSVR model with ELM-Sine and ELM-Sig models, respectively, on the basis of CC as presented in Figure 5(a). Performance enhancements of 54.60% and 63.02% were

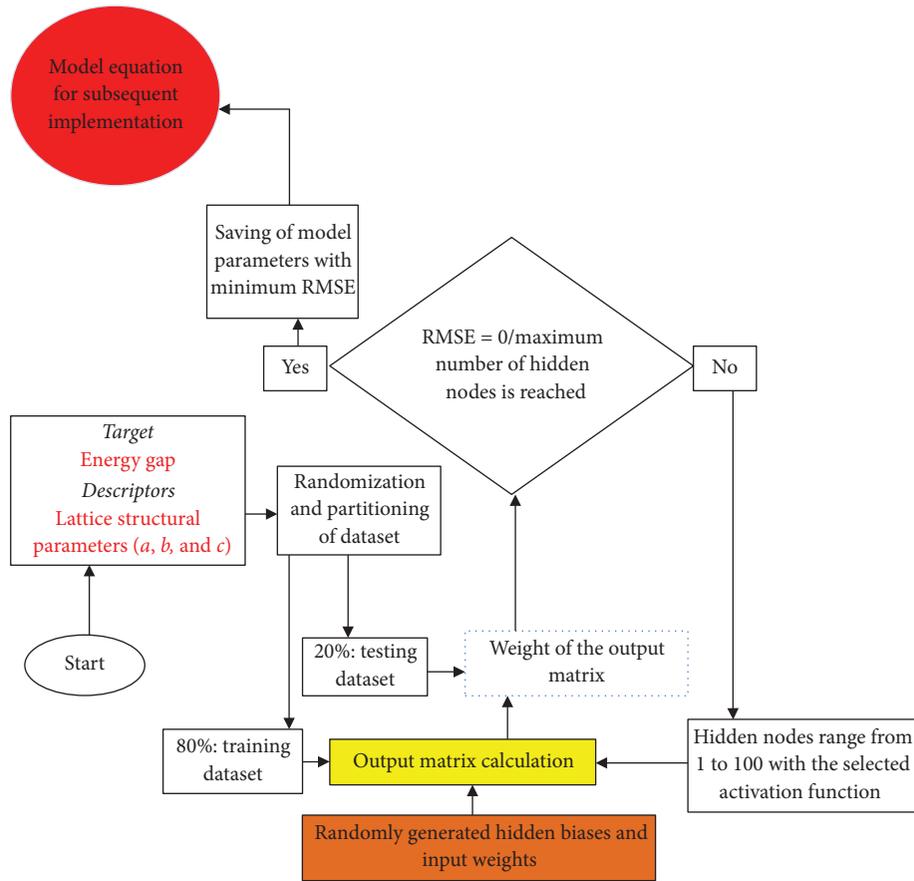


FIGURE 1: Computational flow chat for the developed ELM-based model.

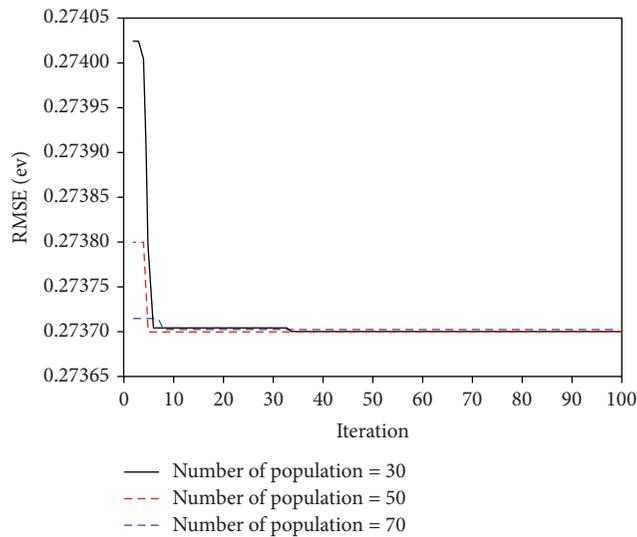


FIGURE 2: Influence of the number of population on the convergence of GSVR model.

obtained during a comparison between the performance of the GSVR and ELM-Sine model on the basis of MAE and RMSE, respectively, for the entire dataset as depicted by Figures 5(b) and 5(c). The actual values obtained for each of

the performance measuring parameters for each model are presented in Table 3. Figure 6 also depicts the cross-plot correlation between the experimental gap energy and the estimated values using each of the developed models.

TABLE 2: Optimum values of the developed models' hyperparameters.

Developed model	Hyperparameters	Optimum value
GSVR model	Epsilon	0.0001
GSVR model	Regularization factor	486.8605
GSVR model	Kernel option	0.002
GSVR model	Number of population	50
GSVR model	Hyperparameter Lambda	E-7
GSVR model	Kernel function	Gaussian
ELM-Sine model	Activation function	Sine function
ELM-Sine model	Number of hidden nodes	4
ELM-Sig model	Activation function	Sigmoid function
ELM-Sig model	Number of hidden nodes	4

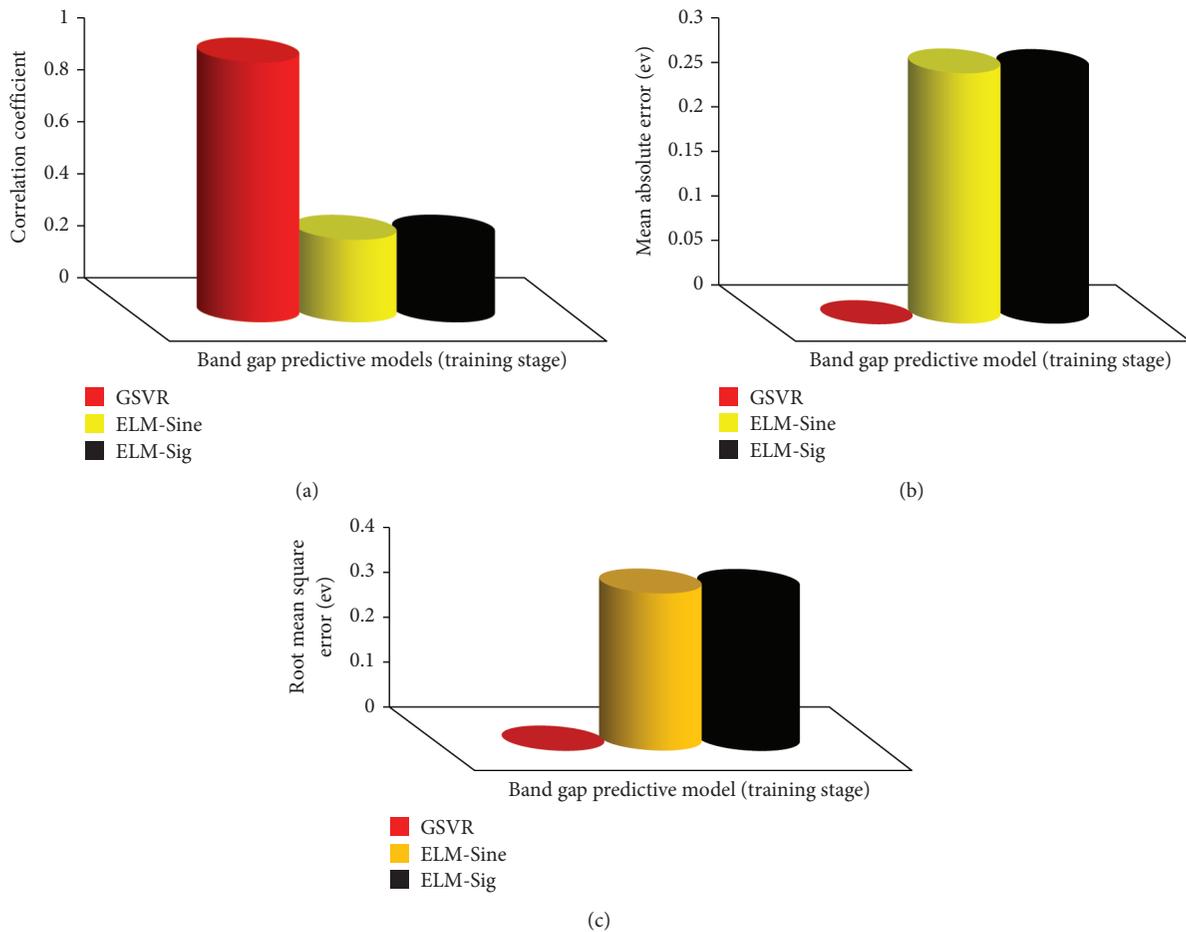


FIGURE 3: Evaluation of the developed models during the training stage of model development.

Alignment of band gap energy estimates of doped barium titanate of GSVR model shows its superiority of other developed models.

4.3. Band Gap Enhancement Effect of Iodine Ion Incorporation on Barium Titanate Semiconductor Using the Developed GSVR Model. The result of iodine ion incorporation on the crystal structure of barium titanate is presented in Figure 7. Substitution of Ti^{4+} ion in the parent barium titanate by I^{5+} iodine ions changes the symmetry and dynamics of the

system. Apart from photo-response enhancement, the efficiency of charge transfer also gains significant improvement. Figure 7 also compares the estimates of the developed GSVR model with the measured values [6]. The estimates of the developed GSVR model align perfectly with the measured values. The observed photocatalytic enhancement can be attributed to the dopant energy states formation near the conduction band edge of the parent barium titanate consequent upon mixing up of iodine 5p states with Ti 3d states which results in shifting oxygen 2p states upward to higher energy in the parent compound.

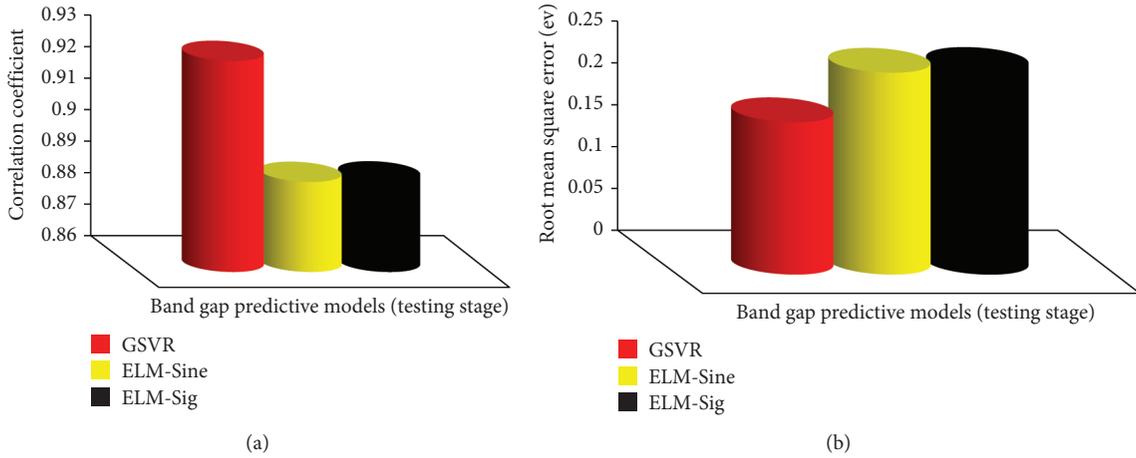


FIGURE 4: Evaluation of the developed models while testing the models.

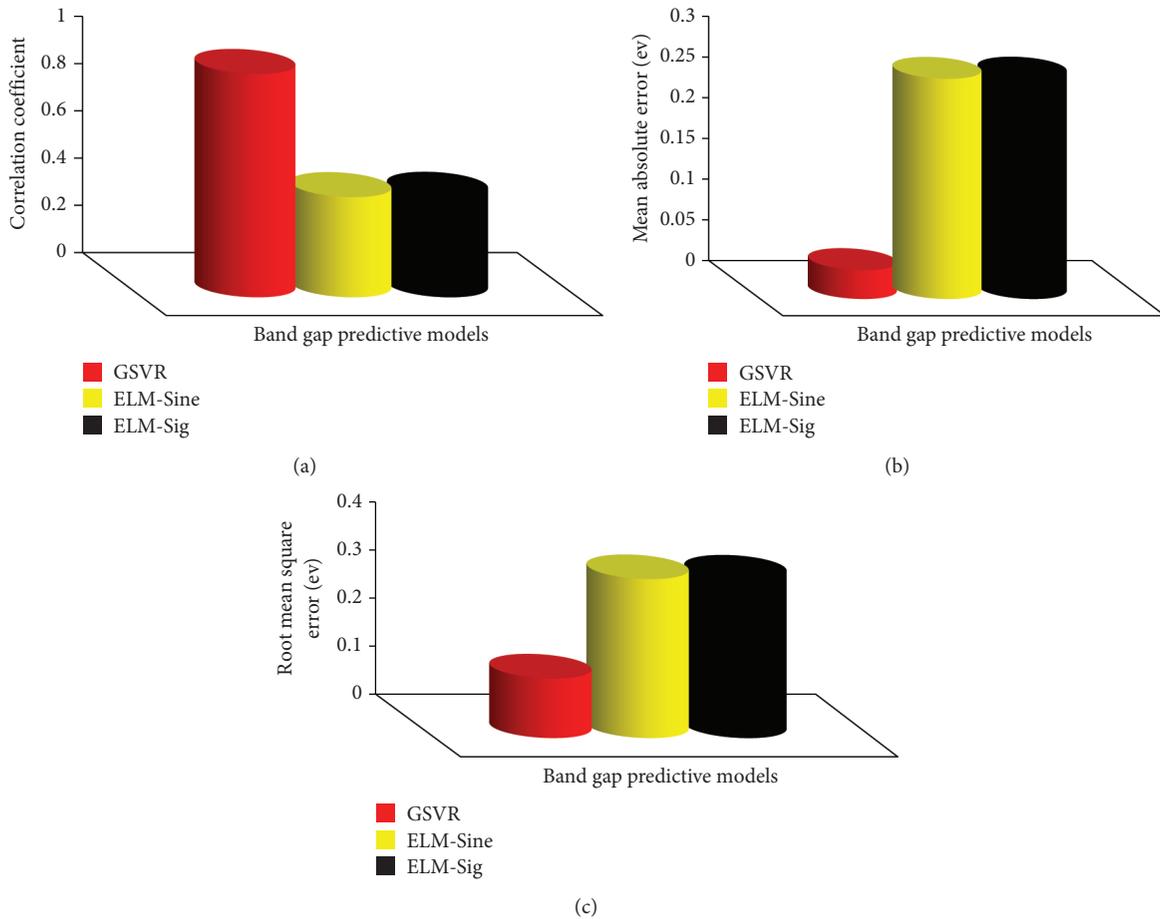


FIGURE 5: Overall evaluation of the developed models.

4.4. Influence of Samarium in Band Gap Improvement of Barium Titanate Semiconductor Using the Developed GSVR Model. Bang gap enhancing capacity of barium titanate doped with samarium ions is presented in Figure 8. Comparison of the measured values [4] with the estimates of the

developed GSVR model shows perfect agreement as presented in the figure.

The observed effect of samarium ions on energy gap improvement can be easily linked to the difference in the ionic state of samarium (Sm^{3+}) as compared to the parent

TABLE 3: Performance evaluation for the developed models.

Developed model		CC	MAE (ev)	RMSE (ev)
GSVR	Training set	1.0000	0.0001	0.0001
	Testing set	0.9266	0.2737	0.1817
	Entire dataset	0.9413	0.0364	0.1224
ELM-Sine	Training set	0.3166	0.2845	0.3498
	Testing set	0.8886	0.2199	0.2420
	Entire dataset	0.4267	0.2716	0.3310
ELM-sig	Training set	0.3167	0.2845	0.3498
	Testing set	0.8888	0.2200	0.2420
	Entire dataset	0.4267	0.2716	0.3310

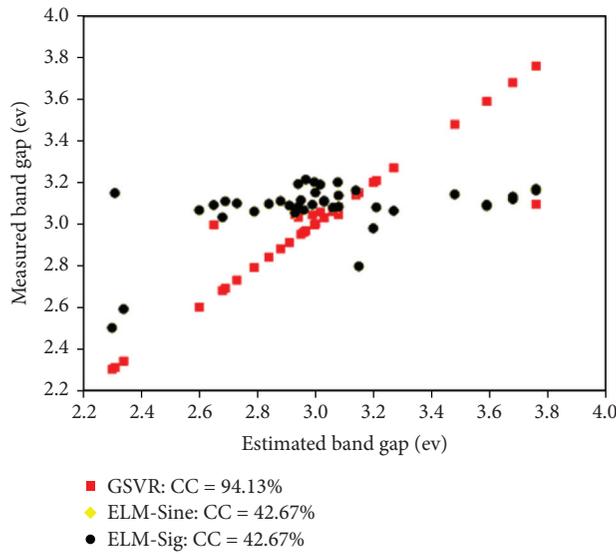


FIGURE 6: Correlation cross-plot for the developed GSVR and ELM-based model.

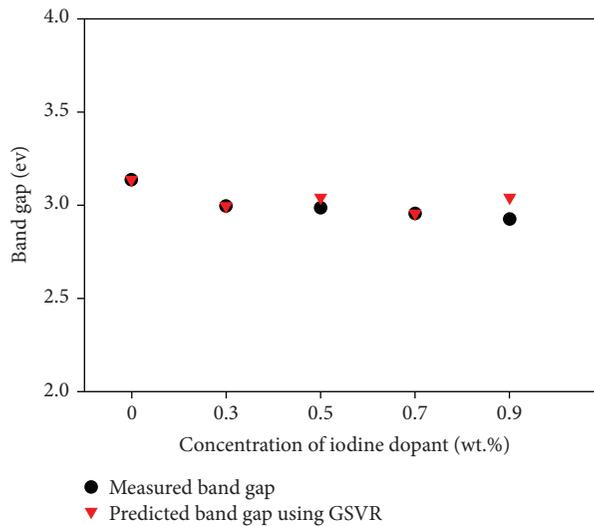


FIGURE 7: Influence of iodine ion incorporation on the band gap of barium titanate for the developed GSVR model.

cation (Ba^{2+}). This ionic state difference changes the dipole region carrier concentration and also leads to the formation of vacancies as well as defects. The developed

GSVR model captures the trend of band gap enhancement effectively due to its hybridization capacity for hyperparameter optimization and the structural risk

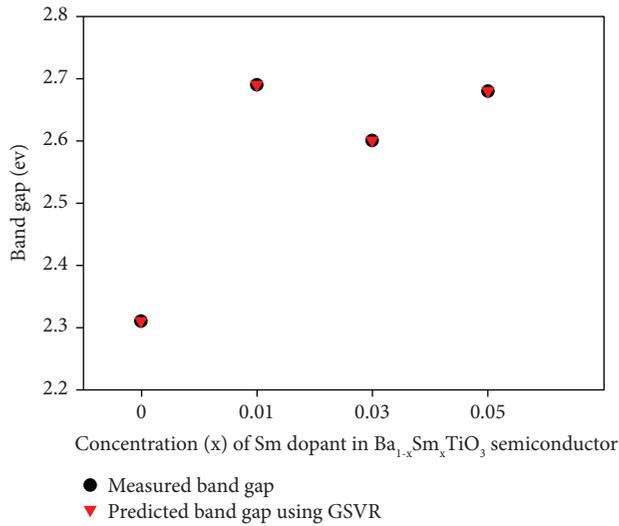


FIGURE 8: Significance of samarium dopants on the band gap of barium titanate for the developed GSVR model.

minimization principle upon which the SVR algorithm was developed.

5. Conclusion

The band gap of barium titanate semiconductor doped with foreign materials is modeled using extreme learning machine and gravitational based support vector regression computational methods. Since dopants incorporation into the host structure of barium titanate creates vacancies, changes the dipole region carrier concentration, and ultimately distorts the lattice structure, crystal lattice parameters of the doped semiconductor serve as the descriptors to the developed models. The developed GSVR model outperforms both ELM-based models using three different performance evaluator parameters. The estimates of the developed GSVR model agree excellently with the experimental band gap while investigating the influence of iodine and samarium dopants on the photocatalytic strength of barium titanate. The obtained performance precision of the developed GSVR model in band gap characterization would ultimately facilitate quick and precise band gap determination of barium titanate with incorporated dopants for photocatalytic applications.

Data Availability

The raw data required to reproduce these findings are available in the cited references in Section 3.1 of the manuscript.

Additional Points

Recommendation. Other machine learning techniques can be explored on this material for the estimation of band gap and other physical properties of barium titanate semiconductor.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] C. Andreea, M. Cernea, E. Corina, G. Aldica, P. Ganea, and R. Trusca, "Lanthanum in fluence on the structure, dielectric properties and luminescence of BaTiO₃ ceramics processed by spark plasma sintering technique," *Journal of Alloys and Compounds*, vol. 706, pp. 538–545, 2017.
- [2] J. Cao, Y. Ji, C. Tian, and Z. Yi, "Synthesis and enhancement of visible light activities of nitrogen-doped BaTiO₃," *Journal of Alloys and Compounds*, vol. 615, pp. 243–248, 2014.
- [3] M. Tian, A. Han, S. Ma, X. Zhu, M. Ye, and X. Chen, "Preparation of Cr-doped BaTiO₃ near infrared reflection pigment powder and its anti-aging performance for acrylonitrile-styrene-acrylate," *Powder Technology*, vol. 378, pp. 182–190, 2021.
- [4] A. Kumari, K. Kumari, F. Ahmed et al., "Influence of Sm doping on structural, ferroelectric, electrical, optical and magnetic properties of BaTiO₃," *Vacuum*, vol. 184, p. 109872, 2021.
- [5] S. M. Yakout, "Journal of solid state chemistry in fluence of Na and Na/Fe doping on the dielectric constant, ferromagnetic and sunlight photocatalytic properties of BaTiO₃ perovskite," *Journal of Solid State Chemistry*, vol. 290, 2020.
- [6] P. M. Nithya and L. G. Devi, "Heavy atom perturbation by the incorporation of iodine ion into BaTiO₃ lattice: reduction of fluorescence and enhancement of rate of interfacial charge transfer process under the visible light irradiation," *Surfaces and Interfaces*, vol. 18, p. 100411, 2020.
- [7] M. Arshad, W. Khan, M. Abushad et al., "Correlation between structure, dielectric and multiferroic properties of lead free Ni modified BaTiO₃ solid solution," *Ceramics International*, vol. 46, no. 17, pp. 27336–27351, 2020.
- [8] B. Choubin, M. Borji, F. S. Hosseini, and A. Mosavi, "Mass wasting susceptibility assessment of snow avalanches using machine learning models," *Scientific Reports*, vol. 10, pp. 1–13, 2020.
- [9] M. Goodarzi and A. A. Dineva, "Groundwater salinity susceptibility mapping using classifier ensemble and bayesian machine learning models," *IEEE Access*, vol. 8, 2020.
- [10] A. Mosavi, F. S. Hosseini, B. Choubin, F. Taromideh, G. Rahi, and A. Dineva, "Susceptibility mapping of soil water erosion using machine learning models," *Water*, vol. 12, pp. 1–17, 1995.
- [11] A. Mosavi, A. Shirzadi, and B. Choubin, "Towards an ensemble machine learning model of random subspace based functional tree classifier for snow avalanche susceptibility mapping," *IEEE Access*, vol. 8, 2020.
- [12] A. Mosavi, F. S. Hosseini, B. Choubin et al., "Susceptibility prediction of groundwater hardness using ensemble machine learning models," *Water*, vol. 12, 2020.
- [13] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, Berlin, Germany, 1995.
- [14] S. M. I. Shamsah and T. O. Owolabi, "Modeling the maximum magnetic entropy change of doped manganite using a grid search-based extreme learning machine and hybrid gravitational search-based support vector regression," *Crystals*, vol. 10, no. 4, p. 310, 2020.
- [15] S. M. I. Shamsah and T. O. Owolabi, "Newtonian mechanics based hybrid machine learning method of characterizing

- energy band gap of doped ZnO semiconductor,” *Chinese Journal of Physics*, vol. 68, no. October, pp. 493–506, 2020.
- [16] T. O. Owolabi and M. A. Gondal, “Quantitative analysis of LIBS spectra using hybrid chemometric models through fusion of extreme learning machines and support vector regression,” *Journal of Intelligent & Fuzzy Systems*, vol. 35, 2018.
- [17] T. O. Owolabi and M. A. Gondal, “Development of hybrid extreme learning machine based chemo-metrics for precise quantitative analysis of LIBS spectra using internal reference pre-processing method,” *Analytica Chimica Acta*, vol. 103, 2018.
- [18] T. A. Oyeohan, I. O. Alade, A. Bagudu, K. O. Sulaiman, S. O. Olatunji, and T. A. Saleh, “Predicting of the refractive index of haemoglobin using the Hybrid GA-SVR approach,” *Computers in Biology and Medicine*, vol. 98, pp. 85–92, 2018.
- [19] I. Lou, Z. Xie, W. K. Ung, and K. M. Mok, “Integrating support vector regression with particle swarm optimization for numerical modeling for algal blooms of freshwater,” *Applied Mathematical Modelling*, vol. 39, no. 19, pp. 5907–5916, 2015.
- [20] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: theory and applications,” *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
- [21] Y. Li, Y. Zeng, Y. Qing, and G.-B. Huang, “Learning local discriminative representations via extreme learning machine for machine fault diagnosis,” *Neurocomputing*, vol. 409, pp. 275–285, 2020.
- [22] Y. Cao, K. Wakil, R. Alyousef et al., “Application of extreme learning machine in behavior of beam to column connections,” *Structures*, vol. 25, no. April, pp. 861–867, 2020.
- [23] Y. Wang, R. Li, and Y. Chen, “Accurate elemental analysis of alloy samples with high repetition rate laser-ablation spark-induced breakdown spectroscopy coupled with particle swarm optimization-extreme learning machine,” *Spectrochimica Acta Part B: Atomic Spectroscopy*, vol. 177, p. 106077, 2021.
- [24] R. Hu, E. Ratner, D. Stewart, K.-M. Björk, and A. Lendasse, “A modified Lanczos Algorithm for fast regularization of extreme learning machines,” *Neurocomputing*, vol. 414, pp. 172–181, 2020.
- [25] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, Berlin, Germany, 1995.
- [26] T. O. Owolabi, “Development of a particle swarm optimization based support vector regression model for titanium dioxide band gap characterization,” *Journal of Semiconductors*, vol. 40, no. 2, 2019.
- [27] T. O. Owolabi, M. A. Suleiman, H. B. Adeyemo, K. O. Akande, J. Alhiyafi, and S. O. Olatunji, “Estimation of minimum ignition energy of explosive chemicals using gravitational search algorithm based support vector regression,” *Journal of Loss Prevention in the Process Industries*, vol. 57, pp. 156–163, 2019.
- [28] A. A. Adewunmi, S. Ismail, T. O. Owolabi, A. S. Sultan, S. O. Olatunji, and Z. Ahmad, “Hybrid intelligent modelling of the viscoelastic moduli of coal fly ash based polymer gel system for water shutoff treatment in oil and gas wells,” *Canadian Journal of Chemical Engineering*, vol. 97, no. 11, 2019.
- [29] M. W. Ahmad, J. Reynolds, and Y. Rezgui, “Predictive modelling for solar thermal energy systems: a comparison of support vector regression, random forest, extra trees and regression trees,” *Journal of Cleaner Production*, vol. 203, pp. 810–821, 2018.
- [30] E. Rashedi, H. Nezamabadi-pour, and S. Saryazdi, “GSA: a gravitational search algorithm,” *Information Sciences*, vol. 179, no. 13, pp. 2232–2248, 2009.
- [31] R. Samala and M. Kotapuri, “Optimal DG sizing and siting in radial system using hybridization of GSA and Firefly algorithms,” *Modelling, Measurement and Control A*, vol. 91, no. 2, pp. 77–82, 2018.
- [32] S. D. Beigvand, H. Abdi, and M. La Scala, “Combined heat and power economic dispatch problem using gravitational search algorithm,” *Electric Power Systems Research*, vol. 133, pp. 160–172, 2016.
- [33] H. Lekeşiz, A. R. Yildiz, and B. S. Yildiz, “Structural design of vehicle components using gravitational search and charged system search algorithms,” *Materials Testing*, vol. 58, no. 1, pp. 79–81, 2016.
- [34] Y. Mohamed Shuaib, M. Surya Kalavathi, and C. Christober Asir Rajan, “Optimal capacitor placement in radial distribution system using Gravitational Search Algorithm,” *International Journal of Electrical Power & Energy Systems*, vol. 64, pp. 384–397, 2015.
- [35] S. Yildiz and A. R. Yildiz, “A comparison of recent meta-heuristic algorithms for crashworthiness optimisation of vehicle thin-walled tubes considering sheet metal forming effects,” *International Journal of Vehicle Design*, vol. 73, no. 1/2/3, pp. 179–188, 2017.
- [36] D. Goswami and S. Chakraborty, “Parametric optimization of ultrasonic machining process using gravitational search and fireworks algorithms,” *Ain Shams Engineering Journal*, vol. 6, no. 1, pp. 315–331, 2015.
- [37] B. Özçelik and I. Ercan, “Investigation of structural, morphological, optical, magnetic and dielectric properties of (1 – X) BaTiO₃/XSr_{0.92}Ca_{0.04}Mg_{0.04}Fe₁₂O₁₉ composites,” *Journal of Magnetism and Magnetic Materials*, vol. 510, 2020.
- [38] L. V. Maneeshya, P. V. Thomas, and K. Joy, “Effects of site substitutions and concentration on the structural, optical and visible photoluminescence properties of Er doped BaTiO₃ thin films prepared by RF magnetron sputtering,” *Optical Materials*, vol. 46, pp. 304–309, 2015.
- [39] L. V. Maneeshya, P. V. Thomas, and K. Joy, “Visible PL emission from erbium doped BaTiO₃ thin films deposited by RF magnetron sputtering,” *Materials Today: Proceedings*, vol. 2, no. 3, pp. 992–996, 2015.
- [40] M. Sharmin, N. Islam, and A. Matin, “Structural modification and evaluation of dielectric and ferromagnetic properties of Ce-modified BiFeO₃-BaTiO₃ ceramics,” *Ceramics International*, vol. 46, pp. 15840–15850, 2020.
- [41] M. Selvaraj, R. Venkatesan, J. Mayandi, and V. Venkatachalapathy, “Influence of tin (IV) doping on structural and optical properties of rhombohedral barium titanate (BaTiO₃),” *Materials Today: proceedings*, vol. 35, no. 3–6, 2021.