ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

Modeling of the Sintered Density in Cu-Al Alloy Using Machine Learning Approaches

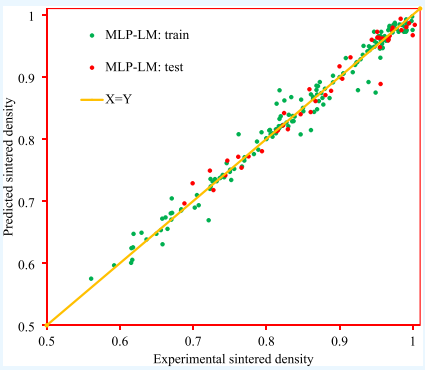
Saleh Asnaashari,* Mohammadhadi Shateri,* Abdolhossein Hemmati-Sarapardeh,* and Shahab S. Band*

Cite This: *ACS Omega* 2023, 8, 28036−28051

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: In powder metallurgy materials, sintered density in Cu-Al alloy plays a critical role in detecting mechanical properties. Experimental measurement of this property is costly and time-consuming. In this study, adaptive boosting decision tree, support vector regression, *k*-nearest neighbors, extreme gradient boosting, and four multilayer perceptron (MLP) models tuned by resilient backpropagation, Levenberg−Marquardt (LM), scaled conjugate gradient, and Bayesian regularization were employed for predicting powder densification through sintering. Yield strength, Young's modulus, volume variation caused by the phase transformation, hardness, liquid volume, liquidus temperature, the solubility ratio among the liquid phase and the solid phase, sintered temperature, solidus temperature, sintered atmosphere, holding time, compaction pressure, particle size, and specific shape factor were regarded as the input parameters of the suggested models. The cross plot, error distribution curve, and cumulative frequency diagram as graphical tools and average percent relative error (APRE), average absolute percent relative error (AAPRE), root mean square error (RMSE), standard deviation (SD), and coefficient of correlation (R) as the statistical evaluations were utilized to estimate the models' accuracy. All of the developed models were compared with preexisting approaches, and the results exhibited that the developed models in the present work are more precise and valid than the existing ones. The designed MLP-LM model was found to be the most precise approach with AAPRE = 1.292%, APRE = −0.032%, SD = 0.020, RMSE = 0.016, and R = 0.989. Lately, outlier detection was applied performing the leverage technique to detect the suspected data points. The outlier detection discovered that few points are located out of the applicability domain of the proposed MLP-LM model.



1. INTRODUCTION

Powder metallurgy (PM), due to its ability to remove the need for secondary operations, is one of the most effective techniques for manufacturing the near-net-shape and complex components, and in some cases, it plays an indispensable role in achieving low-cost parts.[1,2] It also has major advantages like a homogeneous microstructure, lower grain growth, and lower processing temperature.[3] Metallurgical technology includes materials with enhanced properties of strength.[4] Sintering is a crucial stage in the development of raw materials based on powder metallurgy, used in numerous applications.[5] One of key elements in determining the powder metallurgy's mechanical properties is the sintered density.[6−8] The bonding between particles is mainly connected by diffusion. Usually, porosity strongly affects the mechanical properties of the material presents in PM products.[9,10] The literature has shown that the tensile strength decreases by increasing the porosity.[11] Ductility shows more sensitivity to porosity, and also, fatigue resistance is closely correlated with it.[12,13] In the PM method, process parameters and material selection involve inputs from several experts in the field. The process of sintering that results in a material's densification relies on various variables, and the necessary parameters must be adjusted as part of it to organize the required densification, porosity, or connectivity. Construction of standard samples and the design of some testing methods that determine the degree to which densification parameters affect it are costly and time-consuming.[14]

One of the most precise predictive techniques is soft computations that are appropriate to tremendously complicated and multidimensional input/output engineering problems.[15−20] Using these intelligent approaches, a significant number of studies have been accomplished successfully to evaluate many properties in broad metallurgy and material science engineering areas. Design of advanced ultrahigh-strength stainless steels,[21] predicting glass transition temperatures,[22] fracture toughness,[23] coating thickness,[24] compositional optimization,[25] and acoustic properties of tellurite glasses[26] was performed successfully by

Received: November 13, 2022
Accepted: June 21, 2023
Published: July 25, 2023

Table 1. Statistical Features of the Collected Databank in This Study

parameter	minimum	maximum	average	mode	skewness	kurtosis
yield strength (MPa)	15.97	450.00	101.76	29.57	1.73	1.10
Young's modulus (MPa^{-1})	44.00	320.00	139.84	111.92	1.39	0.54
volume variation caused by the phase transformation (100%)	−0.01	0.09	0.01	0.00	1.67	1.17
hardness (HBS)	27.50	333.00	90.74	34.69	1.58	0.71
liquid volume (100%)	0.00	0.28	0.09	0.00	0.49	−1.46
liquidus temperature (°C)	640.00	2613.00	1096.69	1005.00	1.71	3.02
SB/SA	0.00	1.00	0.52	1.00	0.02	−2.00
sintered temperature (°C)	500.00	1400.00	890.69	950.00	0.46	−0.78
solidus temperature (°C)	521.00	3387.00	1028.62	860.00	2.14	8.05
sintered atmosphere (100%)	0.00	39.95	6.72	0.00	1.70	1.18
holding time (h)	0.33	3.00	1.10	1.00	2.43	7.28
compaction pressure (MPa)	100.00	770.00	357.00	374.00	0.07	0.30
particle size (μm)	5.00	87.81	51.29	48.00	−0.40	−1.11
specific shape factor	6.00	11.22	6.94	7.30	2.20	7.69
experimental sintered density	0.56	1.00	0.85	0.96	−0.52	−0.71

using intelligent models. Moreover, some researchers established diverse intelligent techniques for advising on optimal selection of PM materials and processing parameters. These models help in reducing the number of experiments to be carried out, and significant time and cost savings could be achieved.[27−29]

Ma et al.[30] utilized the back-propagation (BP) neural network in synthesizing the nanocomposite WC−18at. %MgO powders to optimize and predict ball milling and processing parameters. Their model determined the correlations between processing parameters and morphological properties of the nanocomposite WC−18at. %MgO. The predictions of the BP neural network presented outstanding agreement with the actual data, and the model precision was proven by error analysis. In 2007, Torkar et al.[31] developed a three-layer ANN with a BP optimizing algorithm procedure to estimate viscosity from the composition of suspension and modeling the behavior of ceramic-paraffin suspensions for low-pressure injection molding (LPIM). The performance of the model shown by comparison between experimental and predicted apparent viscosity curves demonstrated a proper coincidence. Varol and Ozsahin[32] utilized a feed-forward BP-ANN method for predicting the attributes of Al-Cu-Mg flake alloy components synthesized by ball milling. The model inputs were different matrix sizes and different milling times, and the output unit represented specific surface area, apparent density, and flake size. The outcomes derived using the artificial neural network show high similarity among the actual findings and the predicted values. For predicted values, the mean absolute percentage error (MAPE) did not surpass 5.75%. Hence, the prediction ANN model is appropriately operated to predict the Al-Cu-Mg flake alloy. Drndarevic and Reljin[33] built a BP-ANN learning algorithm to precisely define dimensions of PM tools. The used model decreased the dimension deviations of sintered parts with suitably selected process regimes. Ozan et al.[34] used the ANN system for estimating the pore concentration of PM-produced Al-NaCl compacts. Compacting pressure, NaCl content, and NaCl particle size were considered as inputs in the ANN training. The measured and predicted values compared to each other, similarity clarified that the ANN was accurately trained, and statistical values were within acceptable ranges. By using isothermal compression, the hot deformation characteristics of PM Ti-47Al-2Nb-2Cr alloys were studied by Sun et al.[35] After establishing the ANN and Arrhenius-type model, the performance of both models was evaluated with regard to statistical factors. The flow stresses of PM Ti-47Al-2Nb-2Cr alloy were estimated using the BP-ANN model in which the suggested model demonstrated great accuracy and consistency. The coefficient of correlation values for the model and the absolute relative error (ARE) relationship to the prediction were 0.999 and within 5%, respectively. Al-Jabar et al.[36] presented an ANN model to forecast the barium titanate's physical characteristics. They predicted green density, firing shrinkage, porosity, and density of $BaTiO_3$ by using pressing pressure, particle size distribution, pressing rate, sintering temperature, rate of sintering, and soaking times as the input parameters. According to the findings, the mean errors for the green density, shrinkage, density, and porosity were 0.002, 0.04, 0.02, and 0.06, respectively, which exhibited excellent performance of the model. The effect of milling time, volume fraction, and compact pressure on sintered density, hardness, and green density of Al-Al_2O_3 metal matrix composites (MMCs) was predicted using the ANN procedure by Canakci et al.[15] For anticipated values, the mean absolute percentage error (MAPE) did not surpass 5.53%. Consequently, the model is properly utilized to predict the composite mechanical properties accurately. Moreover, machine learning is a helpful tool for the design and optimization of sintering performance.[6]

In the present work, intelligent techniques based on multilayer perceptron (MLP), support vector regression (SVR), *k*-nearest neighbors (KNN), extreme gradient boosting (XGBoost), and adaptive boosting decision tree (AdaBoost-DT) algorithm are implemented to predict the sintered density of metal compact by PM. The goal of the present work is to examine the impact of the changing parameters on sintered density in Cu-9Al alloy. Moreover, various parametric and visual analyses are utilized to assess the efficiency of the developed models. After developing smart models, graphical and statistical error assessments are applied to evaluate the efficiency of the models. In addition, sensitivity analysis is employed to evaluate the consequence of the model's inputs on its output. Afterward, the leverage method is performed for the purpose of identifying the dataset's outlier data points. Finally, the achieved results were compared to models that are published in the literature.

2. MODEL DEVELOPMENT

2.1. Data Collection. The application of a widespread database for modeling is one critical element to determine the

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

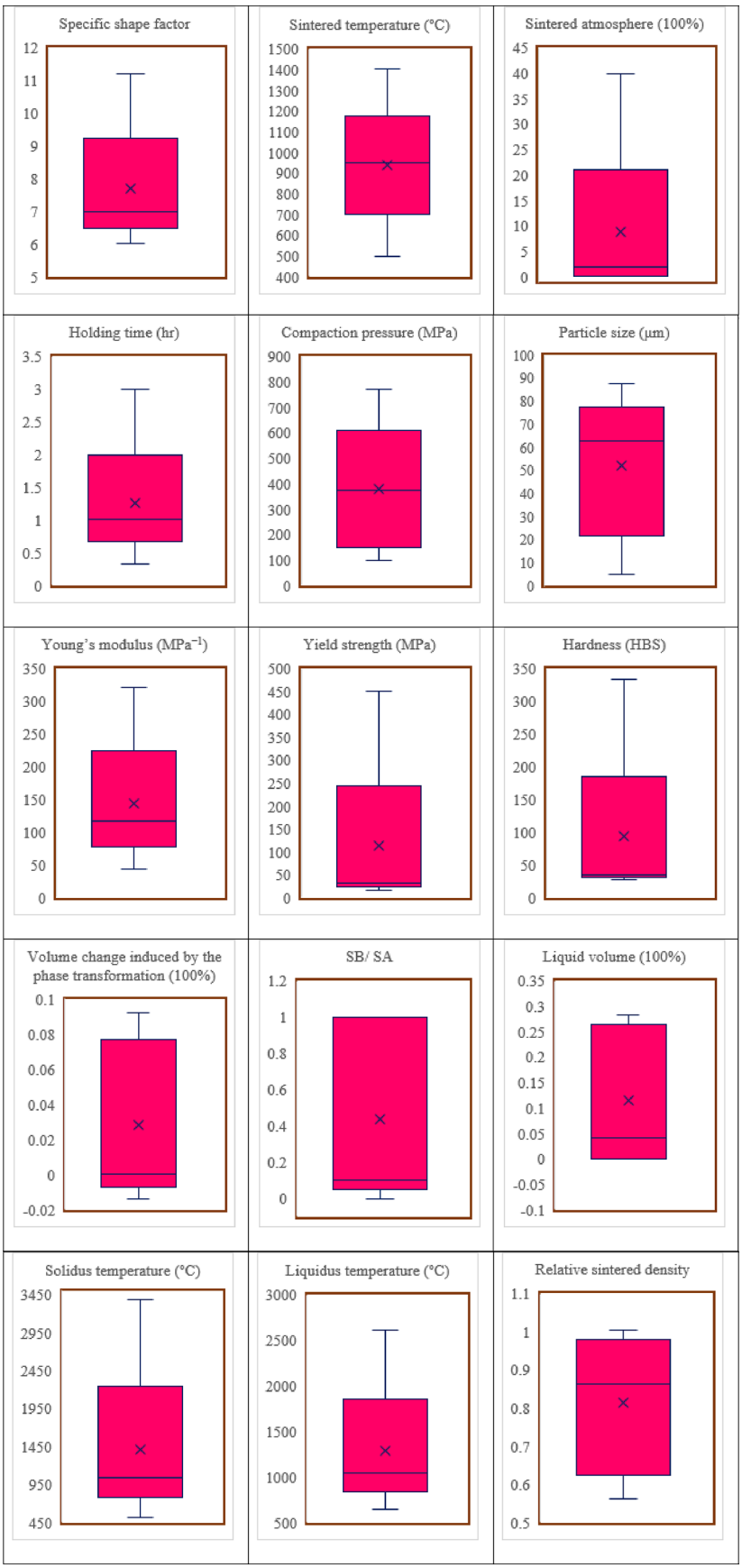


Figure 1. Box-and-whisker graphs for the input and output data points.

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

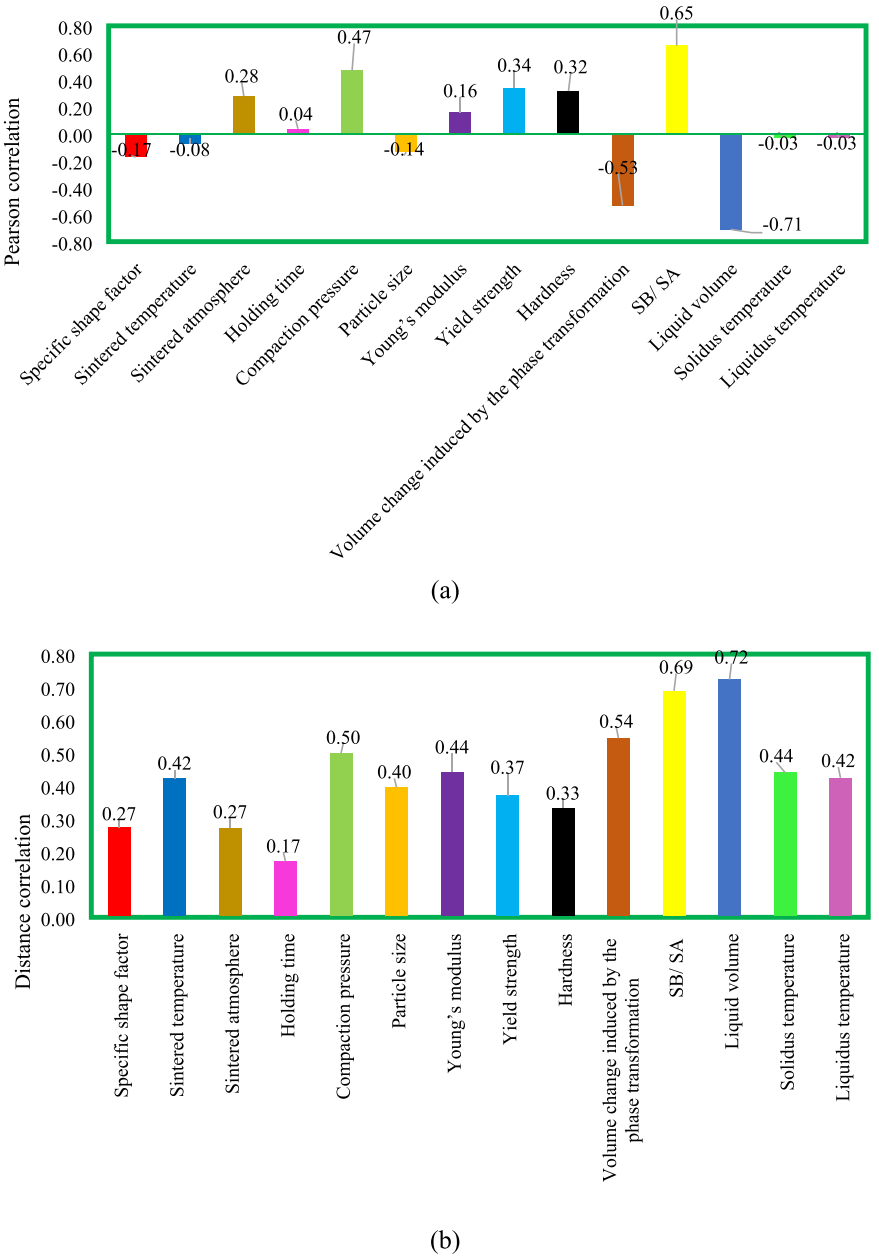


Figure 2. Relative impact of input parameters on output by (a) Pearson correlation and (b) distance correlation.

strength of any modeling study, and these models can be used for predicting output with high certainty. In this study, a large-sized database consisting of 210 experimental sintered density in terms of yield strength, Young's modulus, volume variation caused by the phase transformation, hardness, liquid volume, liquidus temperature, the solubility ratio among the liquid phase and the solid phase (SB/SA), sintered temperature, solidus temperature, sintered atmosphere, holding time, compaction pressure, particle size, and specific shape factor was gathered from open literature,[37−45] as mentioned in ref 6. The statistical assessment of the applied database is shown in Table 1. The distribution of the data is asymmetric in a parameter that deviates from a normal distribution. Skewness defines the asymmetry of a parameter distribution over its mean. Skewness is zero for a normal distribution, and when the skewness is positive, it illustrates that the larger portions of the data are focused in the left-hand side of the function of probability, which means that the amount of low value data is greater than the number of high value data. It is similar for negative values of the skewness. The form of the distribution function can be statistically evaluated as compared to the normal distribution by the kurtosis parameter. Positive kurtosis for unimodal distributions that are symmetric refers to peakedness and heavy tails relative to the normal distribution, while negative kurtosis implies flatness and light tails. A box-and-whisker graph is a reliable plot of describing the dataset on the basis of five factors, namely, minimum, maximum, median (the databank's middle value), first quartile (the first half of the databank's median), and third quartile (the second half of the databank's median). The minimum is at its lowest position, and the maximum point can be seen at the highest peak. Using a flat line in the middle to represent the median, the box is shown from Q1 to Q3 on a graph. The box-and-whisker graphs of inputs and targets are shown in Figure 1.

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

Pearson correlation and distance correlation were utilized to measure linear or nonlinear relationships between the input and output, respectively. With the method of Pearson, values between −1 and 1 are produced. The negative value presents that the output value decreases as this parameter is increased, while the positive value demonstrates that as input increases, the output value intensifies. The distance correlation coefficient is nonnegative, which can take any value in the open unit interval (0 to 1). Pearson correlation and distance correlation for relative impact of input parameters on output are illustrated in Figure 2. In these graphs, it is clear that liquid volume and SB/SA have the greatest effect on sintered density. When a liquid is created during sintering, there are two solubilities that are present: solid phase solubility in the liquid phase and liquid phase solubility in the solid phase. During heating, the uneven diffusivity between the liquid and solid phases results in a change in dimension and shrinkage or swelling happens. Figure 2 represents that volume change induced by the phase transformation and compaction pressure are two other factors that have a major impact on sintered density. During sintering, the lattice type may change, for example, from body-centered cubic to face-centered cubic, and as a result, the volume change due to the phase transformation will be effective on sintered density. Particles slide or rearrange each other when the compaction pressure is low. The primary mechanism for densification at greater pressures is the particle's plastic deformation. High-density compacts can be produced as a result of greater compaction.

2.2. Extreme Gradient Boosting (XGBoost). For minimizing an objective function set that has been used to a training dataset, a collection of different classification and regression trees (CARTs) is used in a tree-based ensemble technique. One of these tree-based models within the gradient boosting decision tree (GBDT) framework is called XGBoost. As shown in Figure 3, the CART's basic structure is made up of

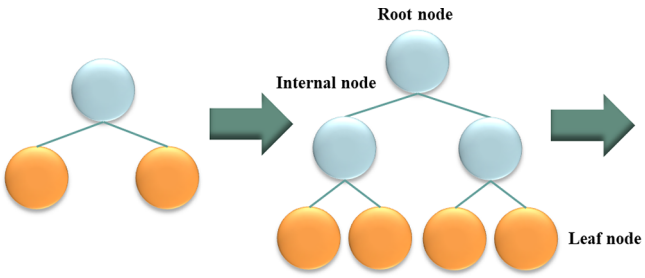


Figure 3. Level-wise tree development in XGBoost.

three different nodes: (a) the main node (root node), (b) internal nodes, and (c) leaf nodes. The root node will be divided into internal nodes by the binary decision-making processes. In this way, the root-located dataset will be divided into different nodes in the internal nodes, and the leaf nodes will serve as the final classes for the final classification. An ensemble of CATRs is introduced and developed using determination of their influence by assigning them a specific weight throughout the training phase, with the goal of generating a powerful set in accordance with the gradient booting model.[46]

To predict the y for a particular dataset, a group of n trees needs to be trained; n and m are counts of instances and features, respectively.

$$\hat{y}_i = \sum_{k=1}^{N} f_k(X_i), f_k \in f$$
$$\text{with} f = \{f(X) = \omega_{q(x)}\}, (q: \mathbb{R}^m \to T, \omega \in \mathbb{R}^T) \quad (1)$$

where using a specified decision rule $q(x)$, a binary leaf index will be created by mapping an example X. The corresponding space of each regression tree is depicted in eqs 1 and 2 by f, f_k shows the kth independent tree, T denotes the number of leaves on the tree, and the weight of the leaf is shown by w, as shown in eqs 2 and 3.

The second stage of the modeling process involves determining tree sets by minimizing an objective function L.

$$L = \sum_{i}^{n} l(\hat{y}_i, y_i) + \sum_{k}^{N} \Omega(f_k)$$
$$\text{with} \Omega(f) = \gamma T + \frac{1}{2}\lambda \|\omega\|^2 \quad (2)$$

where Ω represents the regularization factor that lowers the model's complexity to help prevent overfitting, loss function is demonstrated by l and is fundamentally a differentiable convex function, the minimal loss reduction is represented by γ, which is needed in division of a new leaf, and the coefficient of regulation is denoted by λ. It is important to note that γ and λ facilitate increasing the variance of the model and plummet overfitting.[47] The objective functions for every individual leaf in the approach of boosting are minimized iteratively as follows:

$$L^{(t)} = \sum_{i=1}^{n} \{l(y_i, \hat{y}_i^{(t-1)}) + f_t(X_i)\} + \Omega(f_t) \quad (3)$$

Within the given formula, t represents the tth iteration in the training process. The XGBoost technique greedily provides sufficient space for regression trees, often known as a "greedy algorithm," to noticeably ameliorate the ensemble model. As a result, the objective function is minimized continually to update the output of the XGBoost model:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(X_i) \quad (4)$$

The XGBoost takes use of a shrinkage method that after each level of boosting, a learning factor rate scales freshly increased weights. By minimizing the impact of newly created trees on already formed trees, this prevents overfitting risk.[48]

2.3. *k*-Nearest Neighbors (KNN). One type of algorithm for supervised machine learning that can be applied to both regression and classification issues is the k-nearest neighbor (KNN) method. It is fairly straightforward and easy to implement. In the early 1970s, the KNN was a well-liked non-parametric technique in statistical applications.[49] KNN looks for a set of k samples that, when compared to an unknown sample that is first selected at random, are closest. The k indicates how many neighbors will be used by default and invariant datasets. One of the distance units that can be used to determine how far apart points on a graph are from one another is the Euclidean distance, which is specified in eq 5. As a result, the label of these unidentified samples is established as a typical class by calculating the response variables' average. The performance of the KNN would undoubtedly be significantly impacted by selecting the best value of "K"; thus, by adjusting this parameter throughout a range of values and selecting the optimum value for this hyperparameter, we may determine the optimal value for the

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

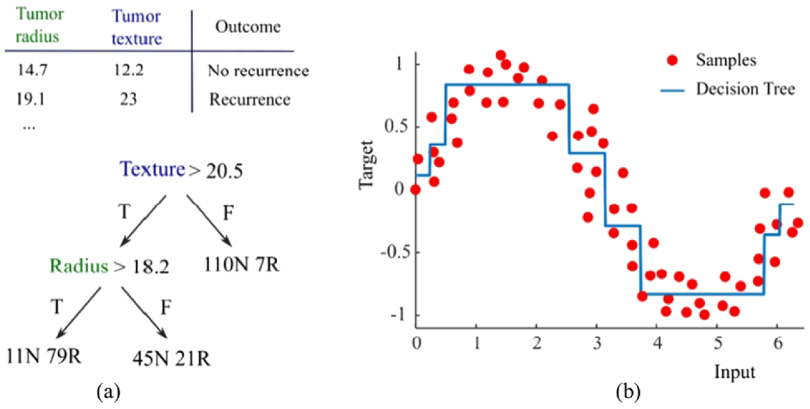


Figure 4. Examples of (a) classification and (b) regression problems where decision trees have been used.[60] Reprinted in part with permission from Nait Amar et al. (2019). Copyright 2019 Elsevier.

parameter "K".[50] The following is how Euclidean distance is calculated (eq 5):

$$D(X, Y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

$$X = (x_1, x_2, ..., x_n) \text{and} Y = (y_1, y_2, ..., y_n) \in \mathbb{R}^n \quad (5)$$

2.4. Support Vector Regression (SVR). Support vector regression (SVR), which has a well specified mathematical model, is frequently utilized for soft calculations, even though support vector machines are a collection of controlled machine learning approaches that can be used for regression and classification.[51] Recently, SVR has aroused the interest of researchers due to its consistency in simulating numerous complex structures. The main theory of SVR has already been published;[52] hence, this paper will only briefly present it. For a dataset $[(x_1, y_1)\ldots. (x_n, y_n)]$ that is given, SVR tries to find a regression function $f(x)$ using the d-dimensional input space $x \in R_d$ and $y \in R$ as the output vector, which depends on the data of the input, to approximate the output as follows:

$$f(x) = w \cdot \phi(x_i) + b \quad (6)$$

where w represents the weight, b shows the bias vectors, and $\phi(x)$ denotes the kernel function. The following minimization technique from Vapnik et al.[53] was described to obtain the weight and bias vectors to the proper values:

$$\text{minimize} \frac{1}{2} w^T w + C \sum_{j=1}^{N} (\zeta_j^- + \zeta_j^+)$$

$$\begin{cases} (w \cdot \phi(x_i) + b) - y_i \le \varepsilon + \zeta_j^- \\ y_i - (w \cdot \phi(x_i) + b) \le \varepsilon + \zeta_j^+ \\ \zeta_j^+, \zeta_j^- \ge 0, i = 1, 2, ..., m \end{cases} \quad (7)$$

Here, T represents the transposed matrix of w, ε denotes the error connivance, C shows the positive regularization parameter representing the variation from ε, and ζ_j^+ and ζ_j^- show the positive parameters representing the higher and lower extra variances, respectively.

By using the Lagrange multipliers, the aforementioned constrained optimization issue is transformed into a dual function. The following is the ultimate solution that results from this action:

$$f(x) = \sum_{j=1}^{n} (a_k - a_k^*) K(x_k, x_l) + b \quad (8)$$

where the kernel function is indicated by $K(x_k, x_l)$, a_k and a_k^* are the multipliers of Lagrange, and b is a bias.

2.5. Adaptive Boosting Decision Tree (AdaBoost-DT). Adaptive boosting-DT is centered on the boosting method and decision tree, one of the effective techniques for improving learning quality. The boosting method developed by Schapire[54] creates a model that is an ensemble method, and to make a high-strength learner, it aims to combine various weak learners/predictors. The process of training a weak learner into a strong learner is done in such a way that emends the prior learners.

2.5.1. Adaptive Boosting (AdaBoost). AdaBoost is a very popular boosting technique that is presented in 1995 by Freund and Schapire[55] for aims of classifying. All the weights are initially set evenly, but the weights of samples that were erroneously categorized are increased in each round such that the new learner is obligated to focus more on the hard examples. The algorithm takes the training data as $\{(x_k, t_k), k = 1,2, \ldots, n\}$ for a classification problem, and the AdaBoost method key steps are represented in Algorithm 1.[56]

Algorithm 1

(steps of AdaBoost Classification).

1. Weight initialization: $w_j = \frac{1}{n}, j = 1,2, \ldots, n$
2. For i=1: N_l (where N_l is the number of learners) do
   a) Set a (weak) learner $Wl_i(x)$ to the training data utilizing data weights w_j
   b) Compute the weighted error rate

$$Err_i = \frac{\sum_{j=1}^{n} w_j I(t_j \ne Wl_i(x))}{\sum_{j=1}^{n} w_j} \text{ where } I(x) = \{0 \text{ if } x = \text{false}, 1 \text{ if } x = \text{true}\}$$

   c) Calculate predictor weights $\beta_i = log\left(\frac{(1 - Err_i)}{Err_i}\right)$
   d) Update data weights $w_j \leftarrow w_j \times exp[\beta_i I(t_j \ne Wl_i(x))], j = 1,2, \ldots, n$
3. Assign $Wl(x) = Sign[\sum_{i=1}^{N_l} Wl_i(x)]$ as output for data test (x)

2.5.2. Decision Tree (DT). This technique is a non-parametric supervised learning method, which can be used to solve both regression and classification issues. The automatic interaction detection (AID) decision tree was created in its initial form by Morgan and Sonquist. The THAID algorithm was developed by Messenger and Mandell as the first classification tree approach.[57] Internal nodes, branches, leaf nodes, and root

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

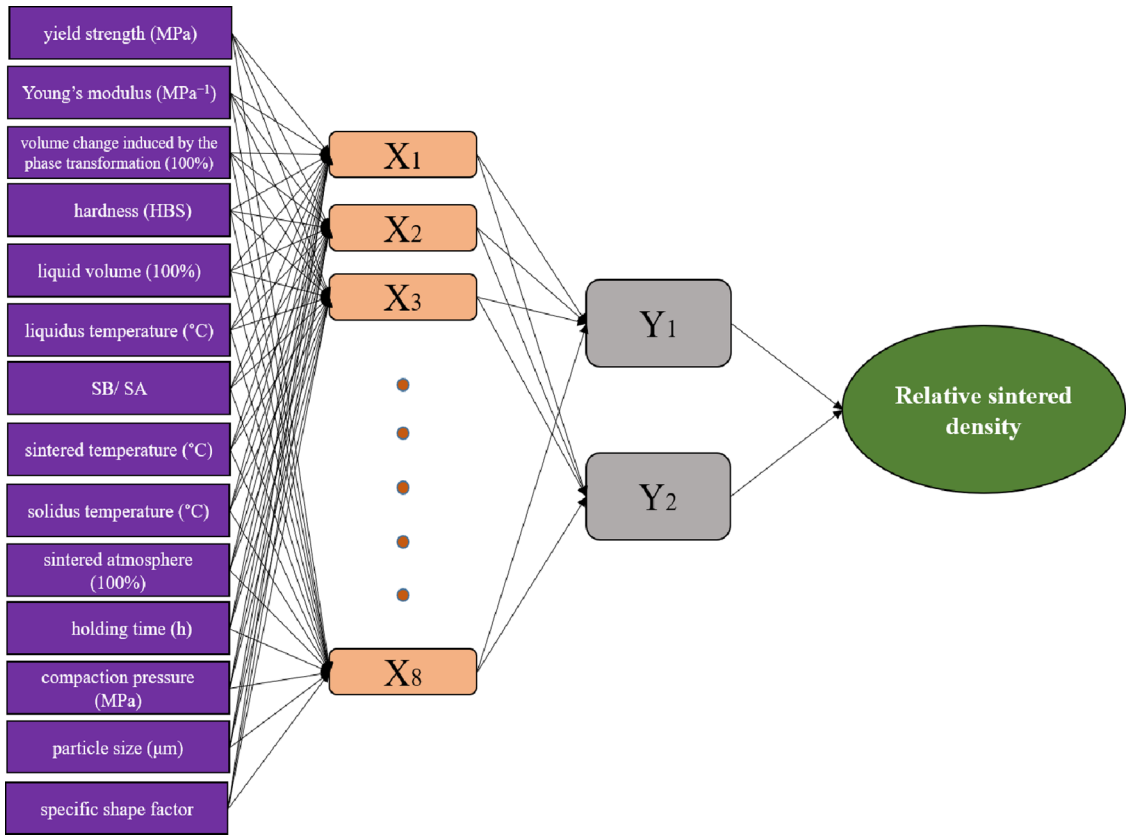


Figure 5. Schematic illustration of the MLP network performed in this work.

nodes are all included in the hierarchical flow diagram that represents this algorithm. The root neuron is the top neuron that reflects the complete sample space and has no income branches. The neuron having a single incoming branch and numerous outgoing edges is known as the test or internal neuron. The other nodes, known as terminal nodes or leaves, represent the end outcomes. Making a decision tree involves three key steps: splitting, stopping, and pruning.[58] Splitting, which can also be used with training instances, is the procedure of breaking the data up into various subsets based on the most prominent attribute. Different factors, such as classification error, Gini index, gain ratio, information gain, and towing, could be taken into account for standard deviation reduction, variance reduction, and classification tree.[59] Figure 4 illustrates a decision tree as an example. As soon as the preset homogeneity is satisfied or the stopping criteria are met, data splitting starts at the root node and moves up to the internal node. By stating the stopping conditions, the complexity of the problem is reduced. The minimal records in a particular node or leaf prior to splitting and the separation between each leaf and root node serve as indicators of this. By adopting this method, overfitting is prevented. Splitting would result in a complicated tree with records that are 100% pure for each node if stopping conditions are not used. In contrast to the test dataset, the training data would be accurately fitted. The model would only be tuned for the best value when stopping criteria are applied. Overfitting is avoided by using the pruning procedure when stopping tactics are ineffective. This approach results in the creation of an entire tree. Then, nodes that contain a smaller validation dataset or information gain are removed to create small trees.

2.6. Multilayer Perceptron (MLP). Artificial neural networks (ANNs), within the field of computational intelligence, are based on the human neuron system similar to the human brain.[61] This approach can determine complex relations between input and output variables. Such networks consist of numerous processing elements (neurons or nodes) that are closely interconnected to one another and links or interconnections (weights). To solve particular issues, nodes or neurons are arranged in different layers, while the weights tie between nodes.[61,62] One of the most popular ANNs is multilayer perceptron (MLP) and includes at least three layers of nodes: an input layer, a hidden layer, and an output layer.[63] The initial layer is related to the input data, and the last one refers to model outputs. Intermediate layers are between these layers, which are named hidden layers, and form the relevance among the model inputs and the eligible outputs.[64] The quantity of input parameters equals to the sum of nodes within the first layer, whereas the neuron number in the last layer is usually equal to one, which is the same as the predicted parameter or property. The process of trial and error is used to match the number of hidden layers and their contained neurons. In several issues, it is acceptable to provide a single hidden layer in an MLP.[65] In general, however, two hidden layers are used for high-complexity structures.[22] The neurons in each hidden layer are linked in their previous and next layers to all the neurons. The sum of the magnitude for each former layer neuron is multiplied for that neuron, at a particular weight, and then, this summation is applied to a bias term to determine the value of each node within the hidden layer or output layer. Afterward, this result is passed through the activation functions. Hidden and output layers can be applied to different activation functions, and some of which are listed below:[66]

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

$$\text{logsig} = \text{sigmoid}: f(x) = \frac{1}{1 + e^{x}} \tag{9}$$

$$\text{linear} = \text{purelin}: f(x) = x \tag{10}$$

$$\text{tansig} = \tanh: f(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} = \frac{2}{1 + e^{-2x}} - 1 \tag{11}$$

In the present study, the most efficient choice was found in a two-hidden-layer MLP with the tansig activation function for the first hidden layer, logsig for the second one, and purelin (linear) transfer function for the layer of the output. Therefore, the output of the model by considering this network can be expressed as follows:[62]

$$\text{output} = \text{purelin}(w_3 \times \log \text{sig}(w_2 \times \tan \text{sig}(w_1 \times x + b_1) + b_2) + b_3 \tag{12}$$

In this equation, w_1 and w_2 are the weight matrixes of the first and second hidden layers, respectively, and w_3 stands for the weight matrix for the output layer. Furthermore, b_1 and b_2 show the bias vectors for the first and second hidden layers, respectively, as b_3 is the output bias. In the performance of the MLP network, choosing the optimization technique for training model is one of the key steps. Instead of using just a single optimization technique that is utilized for most MLP neural network papers, in this paper, four various training algorithms are used. namely, Levenberg−Marquardt (LM), resilient backpropagation (RB), scaled conjugate gradient (SCG), and Bayesian regularization (BR). The resultant MLP approaches established on these algorithms are termed LM-MLP, RB-MLP, SCG-MLP, and BR-MLP, respectively. Figure 5 displays a framework of the MLP network used in this work schematically. Biases and the weights are optimized in these models for the sintered density to minimize the error value. MLP-based models were found to have the best architecture: 14-8-2-1 in all designed networks, in which the first and last numbers represent the number of input and output data, respectively. Moreover, the second and third values indicate how many neurons are present in the first and second hidden layers of the network. Through our modeling, about 80% of the 210 experimental data points and the residual 20% were as the model development and assessing the precision of the established model, respectively. To evade the collection of itemized data points in the problem possible region, randomization was utilized to divide the process into training and testing sets. Table 2 represents the optimum important hyperparameter values for the machine learning approaches applied in the present study.

2.6.1. Resilient Backpropagation Algorithm. Sigmoid and tansig typically utilize as the activation functions in MLP neural networks in the hidden layer. These transfer functions are generally entitled squatting since a finite output range is compacted from an infinite input range. As the large input enters in activation function like sigmoid and tansig, their gradient approaches zero. These can create issues when training the network with the steepest descent since the differential has a small quantity and therefore makes minor adjustments in the biases and weights. To omit the undesirable effects of greatness belonging to partial derivatives, the RB algorithm was proposed.[62]

2.6.2. Levenberg−Marquardt Algorithm. One of the most commonly used methods to optimize the biases and weights in the MLP technique is the algorithm of Levenberg−Marquardt (LM). The most widespread used algorithm was created by Kenneth Levenberg and Donald Marquardt, and it is called Levenberg−Marquardt, presenting a numerical solution for the problem to minimize a nonlinear function.[67−69] Also, this technique is examined on various function approximation difficulties[70] and renowned as the damped least-squares (DLS) technique in which its application is used to resolve nonlinear least-squares issues. In other words, this technique acquires a local minimum that does not inevitably reflect the absolute minimum. In several occasions, this approach can detect the final solution even it is originated from an inaccurate primary guess. The Hessian matrix is not calculated in this technique, and gradient approximation and the Hessian matrix are evaluated as follows if the performance function has the style of sum squares:

$$H = J^{T}J \tag{13}$$

$$g = J^{T}e \tag{14}$$

where e is the network error vector and J is a Jacobian matrix including the first-order derivatives of network errors pertaining to biases and weights. The approximation of the above-mentioned is applied by the LM algorithm in the following Newton-like update:[62]

$$x_{k+1} = x_k - (J^{T}J - \eta I)^{-1}J^{T}e \tag{15}$$

In the above equations, x that stands for the linking η and weights is a constant that lessens after every prosperous step and only rises when the performance function is enlarged by a tentative step. Thus, after each iteration of the algorithm, the performance function value always decreases.

Table 2. Optimum Values of the Hyperparameters for the Proposed Models

	parameters	value
XGBoost	loss	0.6
	max_depth	10
	min_samples_split	2
	min_samples_leaf	1
	n_estimators	70
	subsample	0.7
	max_features	4
KNN	n_neighbors	1
SVR	kernel	radial basis function (rbf)
	C	500
	epsilon	0.001
	gamma	0.3
AdaBoost-DT	loss function	square
	n_estimators	100
	max_depth	15
	min_samples_split	3
	min_samples_leaf	1
	max_features	11
MLP	training function	RB, LM, BR, and SCG
	number of layers	3
	neurons in the first hidden layer	8
	transfer function of the first hidden layer	tansig
	neurons in the second hidden layer	2
	transfer function of the second hidden layer	logsig
	transfer function of the output layer	purelin

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

Table 3. Statistical Error Evaluation of the Developed Models for Determination of Sintered Density Cu-9Al Alloy

	MLP-RB	MLP-LM	MLP-BR	MLP-SCG	AdaBoost-DT	XGBoost	KNN	SVR
				Training set				
AAPRE (%)	1.3760	1.249	1.287	1.335	0.864	2.267	2.849	3.130
APRE (%)	−0.039	−0.054	−0.049	−0.049	−0.093	−1.083	1.656	−0.127
SD	0.020	0.020	0.020	0.021	0.015	0.039	0.048	0.052
RMSE	0.017	0.016	0.016	0.017	0.012	0.028	0.038	0.038
R	0.989	0.990	0.989	0.988	0.994	0.970	0.951	0.940
no. of data	168	168	168	168	168	168	168	168
				Testing set				
AAPRE (%)	1.882	1.463	1.582	1.797	1.864	2.286	2.902	3.266
APRE (%)	0.083	0.055	−0.080	0.037	0.653	0.328	1.660	1.361
SD	0.028	0.020	0.023	0.025	0.028	0.030	0.040	0.043
RMSE	0.022	0.017	0.019	0.020	0.023	0.025	0.033	0.035
R	0.974	0.984	0.983	0.982	0.978	0.971	0.962	0.952
no. of data	42	42	42	42	42	42	42	42
				Total				
AAPRE (%)	1.477	1.292	1.346	1.428	1.064	2.271	2.860	3.157
APRE (%)	−0.014	−0.032	−0.055	−0.032	0.056	−0.801	1.657	0.171
SD	0.022	0.020	0.020	0.022	0.018	0.038	0.046	0.051
RMSE	0.018	0.016	0.017	0.018	0.015	0.028	0.037	0.037
R	0.987	0.989	0.988	0.987	0.991	0.970	0.954	0.942
no. of data	210	210	210	210	210	210	210	210

2.6.3. Bayesian Regularization Algorithm. Bayesian regularization (BR), the renowned training algorithm, will be applied to acquire the weight and bias parameters based on the LM algorithm. This optimization minimizes a grouping of weights and squared errors, and the best arrangement is constructed, which generalizes well. The network's weights are defined as follows for the objective function:[71]

$$F(\omega) = \alpha E_{\omega} + \beta E_{\mathrm{D}} \tag{16}$$

in which the objective function, the sum of model errors, and the sum of squared model weights are, respectively, donated by $F(\omega)$, E_{D}, and, E_{ω}. In the above equation, the coefficients of α and β are objective function parameters that are defined according to the theorem of the Bayes. It is worth noting that the model's weights are regarded as the random variable in BR modeling. The training set and distributions of the weight vectors are also built based on Gaussian distribution. The calculations are transferred to the LM phase for the objective function minimizing; accordingly, consequently, the weight space will be updated. If the stop condition fails to satisfy, then the procedure is repeated and the estimations of α and β are continued.

2.6.4. Scaled Conjugate Gradient Algorithm. In this approach, the weights are set in the sharpest decline direction, which implies that the gradient is the most negative. The performance function decreases more quickly in this direction, but this does not certainly result in the fastest convergence. Employing conjugate gradient leads to a more instant convergence rate compared to the sharpest decline direction and minimizes the deviation in all earlier steps.

$$p_0 = -g_0 \tag{17}$$

In eq 17, the steepest direction of descent for the first iteration is indicated by the symbol $-g_0$ and the search direction, known as the conjugate direction, is indicated by P_0. In this technique, the ideal distance is calculated for optimizing the direction of the current search, as described in eq 18

$$x_{k+1} = x_k + \alpha_k g_k \tag{18}$$

As shown in the following rule, the corresponding search line direction will be defined in conjugating to the preceding search line direction.[66]

$$P_k = -g_k + \beta_k P_{k-1} \tag{19}$$

The conjugate gradient β-determination technique indicates the amount of the conjugate gradient. Apart from the line search technique, it is possible to apply the scaled conjugate gradient (SCG) approach to this purpose that is a cheaper analytical solution than the line search procedure.

3. RESULTS AND DISCUSSION

3.1. Statistical Assessment of the Models. Specific kinds of statistical typicality measures can be computed to analyze the reliability and exactness of each model. The average absolute percent relative error (AAPRE), average percent relative error (APRE), the standard deviation (SD), root mean square error (RMSE), and correlation coefficient (R) are considered as statical parameters for all of the suggested models, which are presented in Table 3.

1. AAPRE:

$$\mathrm{AAPRE} = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{d_{\mathrm{exp},i} - d_{\mathrm{pred},i}}{d_{\mathrm{exp},i}}\right| \tag{20}$$

2. APRE:

$$\mathrm{APRE} = \frac{100}{N}\sum_{i=1}^{N}\left(\frac{d_{\mathrm{exp},i} - d_{\mathrm{pred},i}}{d_{\mathrm{exp},i}}\right) \tag{21}$$

3. SD:

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

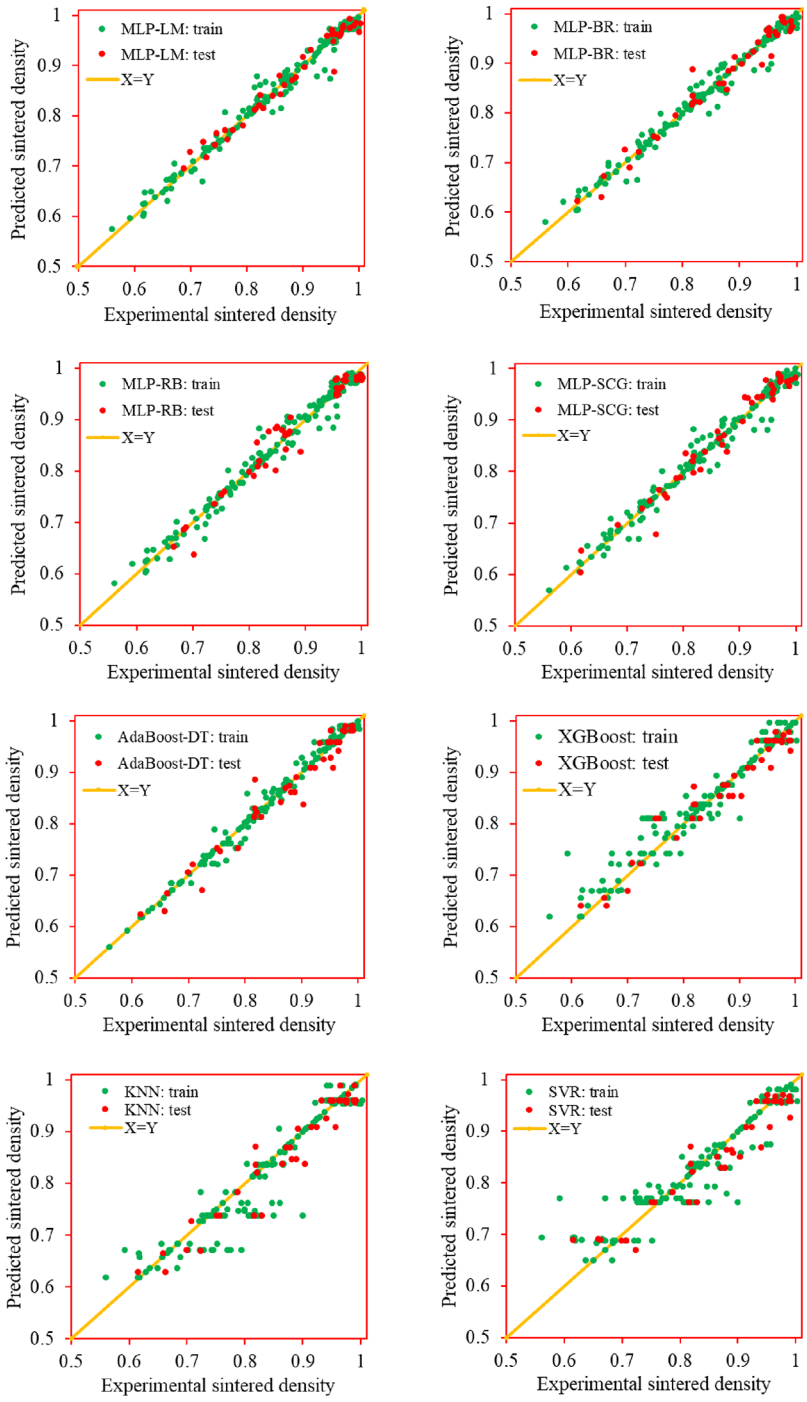


Figure 6. Cross plot for predicted sintered density versus experimental sintered density of train and test sets.

$$\mathrm{SD} = \left(\frac{1}{N-1} \sum_{i=1}^{N} \left(\frac{d_{\mathrm{exp},i} - d_{\mathrm{pred},i}}{d_{\mathrm{exp},i}} \right)^2 \right)^{1/2} \tag{22}$$

4. RMSE:

$$\mathrm{RMSE} = \left(\frac{\sum_{i=1}^{N} (d_{\mathrm{exp},i} - d_{\mathrm{pred},i})^2}{N} \right)^{1/2} \tag{23}$$

5. R:

$$R = \frac{\sum_{i=1}^{N} (d_{\mathrm{exp},i} - \overline{d}_{\mathrm{exp}})(d_{\mathrm{pred},i} - \overline{d}_{\mathrm{pred}})}{\sqrt{\sum_{i=1}^{N} (d_{\mathrm{exp},i} - \overline{d}_{\mathrm{exp}})^2 \sum_{i=1}^{N} (d_{\mathrm{pred},i} - \overline{d}_{\mathrm{pred}})^2}} \tag{24}$$

In eqs 20−24, d, $\overline{d}$, N, and superscripts exp and pred show the sintered density, the average of values, the dataset size, the experimental values, and anticipated values by the published models in this procedure, respectively.

The proposed model validation is confirmed in low values for the RMSE, AAPRE, SD, and APRE parameters in training, testing, and total data points. By taking a look at Table 3, it can be inferred that for all developed models, there is no significant

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

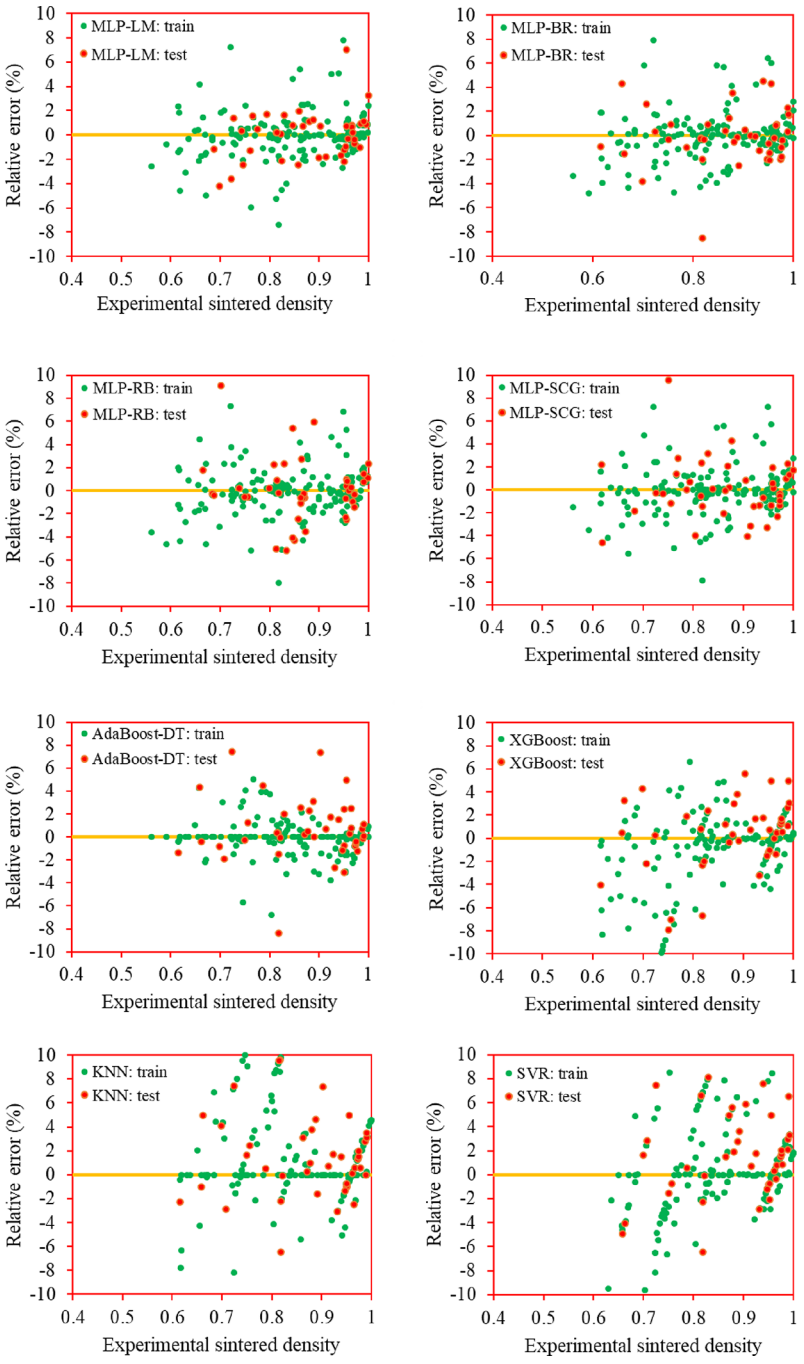


Figure 7. APRE of the predicted sintered density versus experimental sintered density for test and train datasets.

difference between the testing set and the training set except for AdaBoost-DT. Consequently, besides AdaBoost-DT in the present research, the overtraining problem did not happen to all models, which may occur through modeling investigation. Although comparing the statistical parameters of models in Table 3 proves that the AdaBoost-DT strategy has the lowest prediction error in terms of AAPRE, SD, and RMSE compared to the other strategies in the total data points, the MLP-LM is the most exact model as a result of lower values of AAPRE, SD, and RMSE in the testing data points. In contrast, SVR denotes the most prediction error values among the developed models. The statical parameters of the models confirm that BR-MLP and SCG-MLP are valid after LM-MLP. The LM-MLP strategy predicts the sintered density as the output in which the statical parameters consisting of RMSE, AAPRE, SD, APRE, and R are 0.016, 1.249%, 0.020, −0.054%, and 0.990, respectively, for the training set, 0.017, 1.463%, 0.020, 0.055%, and 0.984 for the testing set, and 0.016, 1.292%, 0.020, −0.032%, and 0.989 for the total data points. In consequence, the exactness of the established models in the present study can be ranked as follows: MLP-LM > MLP-BR > MLP-SCG > MLP-RB > AdaBoost-DT > XGBoost > KNN > SVR.

3.2. Graphical Evaluation of the Models. Figure 6 depicts the analysis of graphical error for the developed models as a cross plot of sintered density versus corresponding experimental sintered density for train and test sets. As is evident, tight accumulation of both test and train results around a 45° line reveals that all of the models have perfect accommodation

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

among predicted and experimental measured sintered density. The predictions of the MLP-LM model are accumulated in the closest surrounding area of the unit-slope line, whereas there are no substantial distributed data for every developed approach. In addition, there is no substantial underestimate or overestimate as a result of the scattering pattern.

In Figure 7, the error distribution curves for MLP-LM and other developed models in this study both in testing and training data are plotted. For MLP-LM, most of the scattered data points are gathered surrounding the error line of zero for the entire variety of experimental sintered density. Between the predicted values and the real data, the maximum relative error is 7.85%. This scheme reveals the high consistency degree between the MLP-LM model and the experimental and prediction data over the other ones. In other words, the KNN and the SVR models consist of numerous data that are scattered far from the zero error line.

The other valuable graphical tool is the cumulative frequency of errors in which the MLP-LM, MLP-BR, MLP-SCG, and MLP-RB models as the four best models in this study are represented in Figure 8. As demonstrated in this figure, the

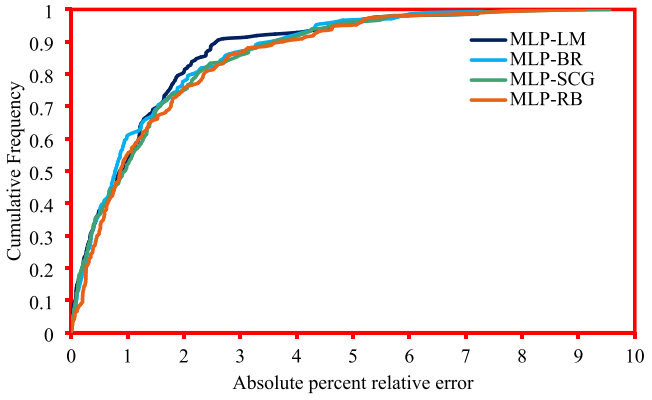


Figure 8. Cumulative frequency curve of MLP-LM, MLP-BR, MLP-SCG, and MLP-RB models.

MLP-LM model presents the highest cumulative frequency of the prediction error. In these models, approximately more than 90% of the predicted sintered density have estimation errors equivalent or fewer than 2.57, 3.52, 3.53, and 3.75% for the MLP-LM, MLP-BR, MLP-SCG, and MLP-RB models, respectively. These results reveal that the model of MLP-LM can be exerted effectively for estimating.

3.3. Comparison between the Developed Models and Models in Literature. Figure 9 exhibits the comparison R values in the published literature with developed techniques in this study. The prediction accuracy increases as the coefficient of correlation approaches 1. The R values for AdaBoost-DT, MLP-LM, MLP-BR, MLP-SCG, MLP-RB, XGBoost, KNN, and SVR are 0.991, 0.989, 0.988, 0.987, 0.987, 0.970, 0.954, and 0.942, respectively. Although the R value for AdaBoost-DT is the maximum, the MLP-LM is the best approach in this work due to the fact that the test data error is low in this model, and the overfitting happened for AdaBoost-DT. The results of most optimization techniques are much more accurate than existing literature techniques of the published literature model[6] in which the coefficient of correlation values were 0.963 for MLP and 0.889 for SVR. Therefore, for both of these models, the corresponding models presented in this study present better

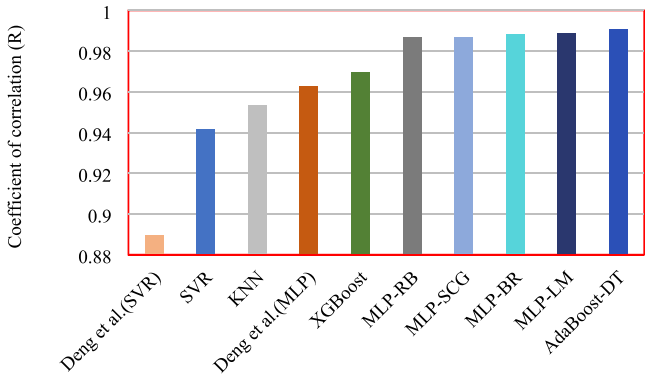


Figure 9. Comparing coefficient of correlation for sintered density models developed (MLP-LM, MLP-RB, MLP-BR, MLP-SCG, and AdaBoost-DT) in this study and the published literature model.

results, which may be related to the issue of modeling design and the algorithm used for modeling.

3.4. Developed MLP-LM Model's Sensitivity Analysis. Sensitivity analysis considers how changes in the input of a model affect its output's value. An accurate way to recognize how the input of a model influences its output is using relevancy factor analysis. For the developed MLP-LM intelligent model in this work, sensitivity analysis was utilized. The relevancy factor computes the influence of each input parameter on the output, and a greater influence of an input on the output is indicated by a higher value of the relevancy factor (r) for that input.[72] Figure 10 represents the impact of input parameters on the relative sintered density. It represents that liquid volume, SB/SA, volume variation caused by the phase transformation, and compaction pressure have the greatest effect on the relative sintered density among the input parameters. Liquid volume and volume variation caused by the phase transformation have a negative effect on the relative sintered density, and increasing these parameters leads to a decrease in relative density. In contrast, SB/SA and compaction pressure have a positive effect on the relative sintered density, and their increase heightens the relative sintered density. Solidus temperature has the least effect among the input parameters with a negative effect.

3.5. Proposed MLP-LM Model's Outlier Detection and Utility Domain. The outlier discovery technique plays an important role to detect data that may be different from the whole data in a dataset. The leverage method is a trustworthy technique for detecting outlier data points. This approach contains the standardized residual values as well as a matrix, the Hat matrix, which is made up of the actual and anticipated values obtained from the developed model.[73]

The Hat indexes are determined based on the Hat matrix (H) with the following formula:

$$H = X(X^t X)^{-1}X^t \tag{25}$$

in which t denotes the transpose matrix and X is described as a ($n \times m$) matrix with n data points and m model input parameters.

Following that, William's plot is created for graphical evaluation of the outliers based on the H matrix calculation. In addition, warning leverage (H^*) is calculated as $\frac{3(x+1)}{n}$, in which both x and n indicate the number of model inputs and data points, respectively. If the majority of the data points fall within the ranges of $0 \leq \text{Hii} \leq H^*$ and $-3 \leq R \leq 3$ (R is the standardized residual), then the model and its predictions are both statistically valid and reliable.[72,73]

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

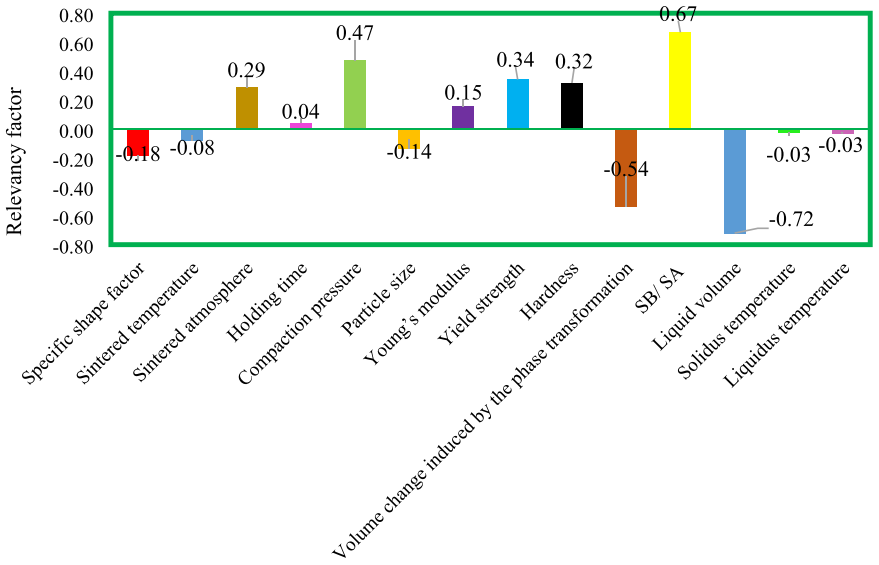


Figure 10. Sensitivity analysis on sintered density behavior of the generated MLP-LM model.

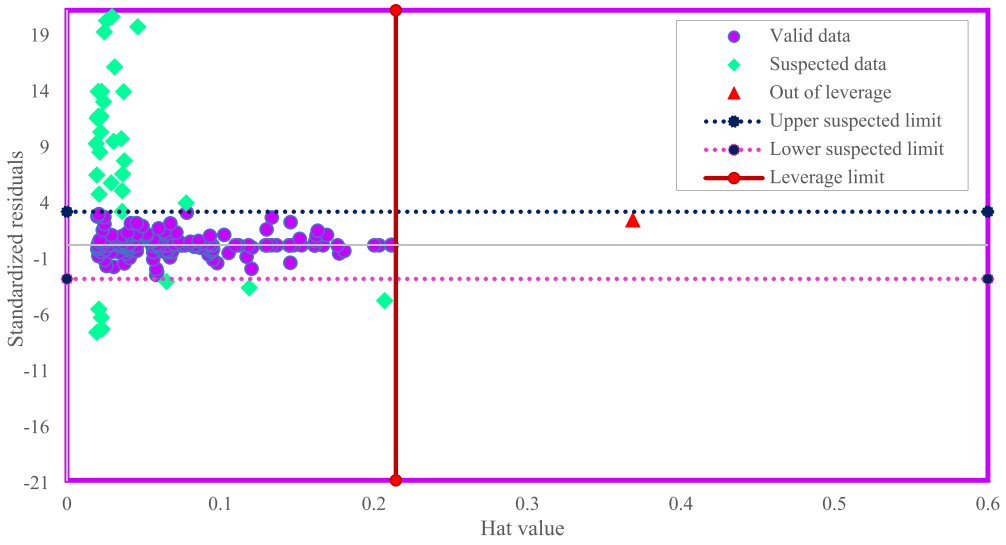


Figure 11. William's plot of the developed AdaBoost-DT model.

The result of the leverage approach is depicted in Figure 11. Majority of the data points exist in the valid zone. Thirty-nine data points (18% of all data) are out of the applicability domain of the model. This confirms that the actual data are reliable, and the model is statistically valid.

4. CONCLUSIONS

In the present paper, eight models were developed depending on the literature data source to estimate the sintered density values in Cu-9Al alloy. For this resolution, 210 data points in terms of yield strength, Young's modulus, volume variation caused by the phase transformation, hardness, liquid volume, liquidus temperature, the solubility ratio among the liquid phase and the solid phase (SB/SA), sintered temperature, solidus temperature, sintered atmosphere, holding time, compaction pressure, particle size, and specific shape factor were collected. Based on the findings of the present study, the following conclusions can be obtained:

1. The models in this work can estimate the sintered density accurately, and MLP-LM has been discovered as the most reliable model.
2. The developed MLP-LM technique predicts the sintered density with high accuracy, including AAPRE values of 1.249% for the training set and 1.463% for the testing set.
3. In the present study, the accuracy of the proposed models is in the following order: MLP-LM > MLP-BR > MLP-SCG > MLP-RB > AdaBoost-DT > XGBoost > KNN > SVR.
4. The developed models based on the soft computing approach were compared to predecessor approaches. The results indicate that the established model's precision was superior to all of those models.
5. With the exception of few data points that were placed in the out of leverage and upper and lower suspected data zones, all of the data points appear to be valid, according to the leverage approach.
6. Liquid volume as one of the input parameters had the most negative influence on the relative sintered density.

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

Moreover, the results indicated that increasing the solubility ratio among the liquid phase and the solid phase (SB/SA) leads to a heightening in the relative sintered density.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.2c07278.

Detailed 210 experimental sintered density in terms of yield strength, Young's modulus, volume variation caused by the phase transformation, hardness, liquid volume, liquidus temperature, the solubility ratio among the liquid phase and the solid phase (SB/SA), sintered temperature, solidus temperature, sintered atmosphere, holding time, compaction pressure, particle size, and specific shape factor as the input parameters; the output of the MLP-LM model (XLSX)

■ AUTHOR INFORMATION

Corresponding Authors

Saleh Asnaashari − *School of Metallurgy and Materials Engineering, University College of Engineering, University of Tehran, Tehran 7761968875, Iran;* orcid.org/0000-0002-8437-6106; Email: esnaashari.saleh@ut.ac.ir

Mohammadhadi Shateri − *Department of System Engineering, École de Technologie Supérieur, Montreal, QC H3C 1K3, Canada*; Email: mohammadhadi.shateri@etsmtl.ca

Abdolhossein Hemmati-Sarapardeh − *Department of Petroleum Engineering, Shahid Bahonar University of Kerman, Kerman 7617117330, Iran;* orcid.org/0000-0002-5889-150X; Email: hemmati@uk.ac.ir

Shahab S. Band − *Future Technology Research Center, College of Future, National Yunlin University of Science and Technology, Douliou, Yunlin 64002, Taiwan, ROC*; Email: shamshirbands@yuntech.edu.tw

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.2c07278

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The respected editor and anonymous reviewers are acknowledged by the authors for taking their time and their constructive comments.

■ REFERENCES

(1) Erdogan, A.; Günen, A.; Gök, M. S.; Zeytin, S. Microstructure and mechanical properties of borided $CoCrFeNiAl_{0.25}Ti_{0.5}$ high entropy alloy produced by powder metallurgy. *Vacuum* **2021**, *183*, No. 109820.

(2) Asnaashari, S.; Ghambari, M. Preparation and characterization of composite WC/Co through rapid omnidirectional compaction. *J. Alloys Compd.* **2021**, *859*, No. 157764.

(3) Kang, B.; Kong, T.; Ryu, H. J.; Hong, S. H. WITHDRAWN: Superior mechanical properties and strengthening mechanisms of lightweight AlxCrNbVMo refractory high-entropy alloys (x = 0, 0.5, 1.0) fabricated by the powder metallurgy process. *J. Mater. Sci. Technol.* **2020**, DOI: 10.1016/j.jmst.2020.08.015.

(4) Stalin, B.; Sudha, G. T.; Ravichandran, M. Optimization of Powder Metallurgy Parameters for AA7072-MoO3 Composites through Taguchi Method. *Mater. Today: Proc.* **2020**, *22*, 2622−2630.

(5) Gómez, S. Y.; Hotza, D. Predicting powder densification during sintering. *J. Eur. Ceram. Soc.* **2018**, *38*, 1736−1741.

(6) Deng, Z.; Yin, H.; Jiang, X.; Zhang, C.; Zhang, K.; Zhang, T.; Xu, B.; Zheng, Q.; Qu, X. Machine leaning aided study of sintered density in Cu-Al alloy. *Comput. Mater. Sci.* **2018**, *155*, 48−54.

(7) Gohar, G. A.; Manzoor, T.; Ahmad, A.; Raza, H.; Farooq, A.; Karim, I.; Iftikhar, W.; Umar, M.; Asad, F. Synthesis and investigate the properties of Cu−Al−Ni alloys with Ag addition using powder metallurgy technique. *J. Alloys Compd.* **2020**, *817*, No. 153281.

(8) Ghasemi, S.; Azadbeh, M.; Mousapour, M.; Mohammadzadeh, A.; Danninger, H.; Salimi, N. The role of pore evolution during supersolidus liquid phase sintering of prealloyed brass powder. *Powder Metall.* **2020**, *63*, 187−196.

(9) Xu, W.; Liu, Z.; Lu, X.; Tian, J.; Chen, G.; Liu, B.; Li, Z.; Qu, X.; Wen, C. Porous Ti-10Mo alloy fabricated by powder metallurgy for promoting bone regeneration. *Sci. China Mater.* **2019**, *62*, 1053−1064.

(10) Shen, W.; Yu, L.; Liu, H.; He, Y.; Zhou, Z.; Zhang, Q. Diffusion welding of powder metallurgy high speed steel by spark plasma sintering. *J. Mater. Process. Technol.* **2020**, *275*, No. 116383.

(11) Cannella, E.; Nielsen, C. V.; Bay, N. Process investigation and mechanical properties of electro sinter forged (ESF) titanium discs. *Int. J. Adv. Des. Manuf. Technol.* **2019**, *104*, 1985−1998.

(12) Laursen, C. M.; DeJong, S. A.; Dickens, S. M.; Exil, A. N.; Susan, D. F.; Carroll, J. D. Relationship between ductility and the porosity of additively manufactured AlSi10Mg. *Mater. Sci. Eng., A* **2020**, *795*, No. 139922.

(13) Biswal, R.; Zhang, X.; Syed, A. K.; Awd, M.; Ding, J.; Walther, F.; Williams, S. Criticality of porosity defects on the fatigue performance of wire + arc additive manufactured titanium alloy. *Int. J. Fatigue* **2019**, *122*, 208−217.

(14) H., Khorsand, M., Arjomandi, H., Abdoos, S., Sadati, Application of artificial neural network for prediction of heat treated sintered steels properties, in: *Defect and Diffusion Forum*; Trans Tech Publ, 2008, pp. 323−328.

(15) Canakci, A.; Varol, T.; Ozsahin, S. Prediction of effect of volume fraction, compact pressure and milling time on properties of Al-Al_2O_3 MMCs using neural networks. *Met. Mater. Int.* **2013**, *19*, 519−526.

(16) Radha, P.; Chandrasekaran, G.; Selvakumar, N. Simplifying the powder metallurgy manufacturing process using soft computing tools. *Appl. Soft Comput.* **2015**, *27*, 191−204.

(17) Patowari, P. K.; Saha, P.; Mishra, P. K. Artificial neural network model in surface modification by EDM using tungsten−copper powder metallurgy sintered electrodes. *Int. J. Adv. Des. Manuf. Technol.* **2010**, *51*, 627−638.

(18) Ershadnia, R.; Amooie, M. A.; Shams, R.; Hajirezaie, S.; Liu, Y.; Jamshidi, S.; Soltanian, M. R. Non-Newtonian fluid flow dynamics in rotating annular media: Physics-based and data-driven modeling. *J. Pet. Sci. Eng.* **2020**, *185*, No. 106641.

(19) Anitescu, C.; Atroshchenko, E.; Alajlan, N.; Rabczuk, T. Artificial neural network methods for the solution of second order boundary value problems. *Comput. Mater. Contin.* **2019**, *59*, 345−359.

(20) Samaniego, E.; Anitescu, C.; Goswami, S.; Nguyen-Thanh, V. M.; Guo, H.; Hamdia, K.; Zhuang, X.; Rabczuk, T. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. *Comput. Methods Appl. Mech. Eng.* **2020**, *362*, No. 112790.

(21) Shen, C.; Wang, C.; Wei, X.; Li, Y.; van der Zwaag, S.; Xu, W. Physical metallurgy-guided machine learning and artificial intelligent design of ultrahigh-strength stainless steel. *Acta Mater.* **2019**, *179*, 201−214.

(22) Cassar, D. R.; de Carvalho, A. C. P. L. F.; Zanotto, E. D. Predicting glass transition temperatures using neural networks. *Acta Mater.* **2018**, *159*, 249−256.

(23) Liu, G.; Jia, L.; Kong, B.; Guan, K.; Zhang, H. Artificial neural network application to study quantitative relationship between silicide and fracture toughness of Nb-Si alloys. *Mater. Des.* **2017**, *129*, 210−218.

(24) Varol, T.; Canakci, A.; Ozsahin, S.; Erdemir, F.; Ozkaya, S. Artificial neural network-based prediction technique for coating

thickness in Fe-Al coatings fabricated by mechanical milling. *Part. Sci. Technol.* **2018**, *36*, 742−750.
(25) Möller, J. J.; Körner, W.; Krugel, G.; Urban, D. F.; Elsässer, C. Compositional optimization of hard-magnetic phases with machine-learning models. *Acta Mater.* **2018**, *153*, 53−61.
(26) Gaafar, M. S.; Abdeen, M. A. M.; Marzouk, S. Y. Structural investigation and simulation of acoustic properties of some tellurite glasses using artificial intelligence technique. *J. Alloys Compd.* **2011**, *509*, 3566−3575.
(27) Arif, S.; Alam, M. T.; Ansari, A. H.; Shaikh, M. B. N.; Siddiqui, M. A. Analysis of tribological behaviour of zirconia reinforced Al-SiC hybrid composites using statistical and artificial neural network technique. *Mater. Res. Express* **2018**, *5*, No. 056506.
(28) Taskin, M.; Dikbas, H.; Caligulu, U. Artificial neural network (ann) approach to prediction of diffusion bonding behavior (shear strength) of ni-ti alloys manufactured by powder metalurgy method. *Math. Comput. Appl.* **2008**, *13*, 183−191.
(29) Varol, T.; Canakci, A.; Ozsahin, S. Prediction of effect of reinforcement content, flake size and flake time on the density and hardness of flake AA2024-SiC nanocomposites using neural networks. *J. Alloys Compd.* **2018**, *739*, 1005−1014.
(30) Ma, J.; Zhu, S. G.; Wu, C. X.; Zhang, M. L. Application of back-propagation neural network technique to high-energy planetary ball milling process for synthesizing nanocomposite WC−MgO powders. *Mater. Des.* **2009**, *30*, 2867−2874.
(31) Torkar, D.; Novak, S.; Novak, F. Apparent viscosity prediction of alumina−paraffin suspensions using artificial neural networks. *J. Mater. Process. Technol.* **2008**, *203*, 208−215.
(32) Varol, T.; Ozsahin, S. Artificial neural network analysis of the effect of matrix size and milling time on the properties of flake Al-Cu-Mg alloy particles synthesized by ball milling. *Part. Sci. Technol.* **2019**, *37*, 381−390.
(33) Drndarevic, D.; Reljin, B. Accuracy modelling of powder metallurgy process using backpropagation neural networks. *Powder Metall.* **2000**, *43*, 25−29.
(34) Ozan, S.; Taskin, M.; Kolukisa, S.; Ozerdem, M. S. Application of ANN in the prediction of the pore concentration of aluminum metal foams manufactured by powder metallurgy methods. *Int. J. Adv. Des. Manuf. Technol.* **2008**, *39*, 251−256.
(35) Sun, Y.; Hu, L.; Ren, J. Modeling the constitutive relationship of powder metallurgy Ti-47Al-2Nb-2Cr alloy during hot deformation. *J. Mater. Eng. Perform.* **2015**, *24*, 1313−1321.
(36) Al-Jabar, A. J. A.; Al-Dujaili, M. A. A.; Al-Hydary, I. A. D. Prediction of the physical properties of barium titanates using an artificial neural network. *Appl. Phys. A: Mater. Sci. Process.* **2017**, *123*, 274.
(37) Hwang, K. S.; Huang, H. S. The liquid phase sintering of molybdenum with Ni and Cu additions. *Mater. Chem. Phys.* **2001**, *67*, 92−100.
(38) Showaiter, N.; Youseffi, M. Compaction, sintering and mechanical properties of elemental 6061 Al powder with and without sintering aids. *Mater. Des.* **2008**, *29*, 752−762.
(39) FENG, Y.; LI, Y.-m.; HE, H.; ZENG, Z.-y. Sintering expansion behavior and expansion mechanism of Cu-Ni-A_l powder alloy. *J. Cent. South Univ.* **2010**, *1*, 108−113.
(40) Fathy, A.; Omyma, E.-K.; Mohammed, M. M. M. Effect of iron addition on microstructure, mechanical and magnetic properties of Al-matrix composite produced by powder metallurgy route. *Trans. Nonferrous Met. Soc. China* **2015**, *25*, 46−53.
(41) Hao, H.; Ye, S.; Yu, K.; Chen, P.; Gu, R.; Yu, P. The role of alloying elements on the sintering of Cu. *J. Alloys Compd.* **2016**, *684*, 91−97.
(42) Hao, H.; Mo, W.; Lv, Y.; Ye, S.; Gu, R.; Yu, P. The effect of trace amount of Ti and W on the powder metallurgy process of Cu. *J. Alloys Compd.* **2016**, *660*, 204−207.
(43) Jingyuan, Y.; Jianzhong, W.; Qiang, L.; Jian, S.; Jianming, C.; Xudong, S. Effect of Zn on microstructures and properties of Mg-Zn alloys prepared by powder metallurgy method. *Rare Met. Mater. Eng.* **2016**, *45*, 2757−2762.
(44) Xu, Z.; Hodgson, M. A.; Cao, P. Effects of mechanical milling and sintering temperature on the densification, microstructure and tensile properties of the Fe−Mn−Si powder compacts. *J. Mater. Sci. Technol.* **2016**, *32*, 1161−1170.
(45) Benavides, P. A.; Soto, B.; Palma, R. H. Liquid phase sintering of mechanically alloyed Mo-Cu powders. *Mater. Sci. Eng., A* **2017**, *701*, 237−244.
(46) T., Chen, C., Guestrin, Xgboost: A scalable tree boosting system, in: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*; Association for Computing Machinery 2016, pp. 785−794.
(47) Zhang, J.; Sun, Y.; Shang, L.; Feng, Q.; Gong, L.; Wu, K. A unified intelligent model for estimating the (gas+ n-alkane) interfacial tension based on the eXtreme gradient boosting (XGBoost) trees. *Fuel* **2020**, *282*, 118783.
(48) V.A., Dev, M.R., Eden, Gradient boosted decision trees for lithology classification, in: *Computer aided chemical engineering*; Elsevier, 2019, pp. 113−118.
(49) Altman, N. S. An introduction to kernel and nearest-neighbor nonparametric regression. *Am. Stat.* **1992**, *46*, 175−185.
(50) Thanh Noi, P.; Kappas, M. J. S. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors* **2018**, *18*, 18.
(51) Smola, A. J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199−222.
(52) Schölkopf, B.; Smola, A. J.; Williamson, R. C.; Bartlett, P. L. New support vector algorithms. *Neural Comput.* **2000**, *12*, 1207−1245.
(53) Vapnik, V.; Golowich, S.; Smola, A. Support vector method for function approximation, regression estimation and signal processing. *Adv. Neural Inf. Process. Syst.* **1996**, *9*.
(54) Schapire, R. E. The strength of weak learnability. *Mach. Learn.* **1990**, *5*, 197−227.
(55) Freund, Y.; Schapire, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119−139.
(56) J., Friedman, T., Hastie, R., Tibshirani, *The elements of statistical learning*. vol. *1* Springer series in statistics, New York, (2001).
(57) Loh, W. Y. Fifty years of classification and regression trees. *Int. Stat. Rev.* **2014**, *82*, 329−348.
(58) Song, Y.-Y.; Ying, L. Decision tree methods: applications for classification and prediction. *Shanghai Arch. Psychiatry* **2015**, *27*, 130−135.
(59) Patel, N.; Upadhyay, S. Study of various decision tree pruning methods with their empirical comparison in WEKA. *Int. J. Comput. Appl.* **2012**, *60*, 20−25.
(60) Nait Amar, M.; Shateri, M.; Hemmati-Sarapardeh, A.; Alamatsaz, A. Modeling oil-brine interfacial tension at high pressure and high salinity conditions. *J. Pet. Sci. Eng.* **2019**, *183*, No. 106413.
(61) Karkevandi-Talkhooncheh, A.; Rostami, A.; Hemmati-Sarapardeh, A.; Ahmadi, M.; Husein, M. M.; Dabir, B. Modeling minimum miscibility pressure during pure and impure CO2 flooding using hybrid of radial basis function neural network and evolutionary techniques. *Fuel* **2018**, *220*, 270−282.
(62) Hemmati-Sarapardeh, A.; Varamesh, A.; Husein, M. M.; Karan, K. On the evaluation of the viscosity of nanofluid systems: Modeling and data assessment. *Renewable Sustainable Energy Rev.* **2018**, *81*, 313−329.
(63) Hemmati-Sarapardeh, A.; Ameli, F.; Varamesh, A.; Shamshirband, S.; Mohammadi, A. H.; Dabir, B. Toward generalized models for estimating molecular weights and acentric factors of pure chemical compounds. *Int. J. Hydrogen Energy* **2018**, *43*, 2699−2717.
(64) Allahkarami, E.; Salmani Nuri, O.; Abdollahzadeh, A.; Rezai, B.; Maghsoudi, B. Improving estimation accuracy of metallurgical performance of industrial flotation process by using hybrid genetic algorithm−artificial neural network (GA-ANN). *Physicochem. Probl. Miner. Process.* **2017**, *53*.
(65) Parvizi, S.; Hafizpour, H. R.; Sadrnezhaad, S. K.; Akhondzadeh, A.; Abbasi Gharacheh, M. Neural network prediction of mechanical

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051

properties of porous NiTi shape memory alloy. *Powder Metall.* **2011**, *54*, 450−454.
(66) Ameli, F.; Hemmati-Sarapardeh, A.; Schaffie, M.; Husein, M. M.; Shamshirband, S. Modeling interfacial tension in N2/n-alkane systems using corresponding state theory: Application to gas injection processes. *Fuel* **2018**, *222*, 779−791.
(67) H., Yu, B.M., Wilamowski, Levenberg−marquardt training, in: *Intelligent systems*; CRC Press, 2018, pp. 12-11-12-16.
(68) A., Ranganathan, *The Levenberg−Marquardt algorithm*; Tutoral on LM Algorithm, (2004).
(69) J.J., Moré, The Levenberg-Marquardt algorithm: implementation and theory, in: *Numerical analysis*, Springer, 1978, pp. 105−116.
(70) M.-H., Fun, M.T., Hagan, Levenberg-Marquardt training for modular networks, in: *Proceedings of International Conference on Neural Networks (ICNN'96)*; IEEE, 1996, pp. 468−473.
(71) Z., Yue, Z., Songzheng, L., Tianshi, Bayesian regularization BP Neural Network model for predicting oil-gas drilling cost, in: *2011 International Conference on Business Management and Electronic Information*; IEEE, 2011, pp. 483−487.
(72) Amiri-Ramsheh, B.; Safaei-Farouji, M.; Larestani, A.; Zabihi, R.; Hemmati-Sarapardeh, A. Modeling of wax disappearance temperature (WDT) using soft computing approaches: Tree-based models and hybrid models. *J. Pet. Sci. Eng.* **2022**, *208*, 109774.
(73) A., Hemmat-Sarapardeh, A., Larestani, N.A., Menad, S., Hajirezaie, *Applications of artificial intelligence techniques in the petroleum industry*; Gulf Professional Publishing, 2020.

https://doi.org/10.1021/acsomega.2c07278
ACS Omega 2023, 8, 28036−28051