

A Hybrid Material Generation Algorithm with Probabilistic Neural Networks for Solving Classification Problems

Mohammad Wedyan¹, Omar Alshaweesh², Enas Ramadan³
Ryan Alturki⁴, Foziah Gazzawe⁵, Mohammed J. Alghamdi⁶

Department of Autonomous Systems, Al-Balqa Applied University, Al-Salt, Jordan¹
Department of Software Engineering, AL-Hussein Bin Talal University, Ma'an, Jordan²
Department of Computer Science, Al-Balqa Applied University, Al-Salt, Jordan³
Department of Information Science, College of Computer and Information Systems^{4, 5, 6}
Umm Al-Qura University, Makkah, Saudi Arabia^{4, 5, 6}

Abstract—Classification is based on machine learning, in which each element in a set of data is classified into one of a predetermined set of groups. In data mining, an artificial neural network (ANN) is the most significant methodology because of the exact results obtained through this algorithm and applied in solving many classification problems. ANN consists of a group of types of feed-forward networks, feed-back network, RBF networks, and the probabilistic neural networks (PNN). For classification issues, the PNN is frequently utilized. The primary goals of this research are to fine-tune the weights of neural networks to enhance the classification accuracy. To accomplish this goal, the Material Generation Algorithm (MGA) was investigated with PNN in a hybrid model. Newly, the hybridization of algorithms is ubiquitous and it has led to the development of unique procedures that outperform those that use a single algorithm. Several distinct classification tasks are used to test the efficiency of the suggested (MGA-PNN) approach. The MGA algorithm's efficiency is evaluated using the PNN training outcomes generated, and its outcomes are compared to that of other optimization strategies. By 11 benchmark datasets, the suggested algorithm's performance in terms of classification accuracy is evaluated. The outcomes display that the MGA outperforms the biogeography based optimization, firefly method in terms of classification accuracy.

Keywords—Artificial neural network (ANN); material generation algorithm (MGA); classification; probabilistic neural networks (PNN)

I. INTRODUCTION

Recently, our ability to collect data has greatly improved [1]. Millions of databases have been utilized in a variety of applications, including marketing campaigns, company management, scientific endeavors, and several others [2]. The availability of sophisticated and affordable database systems has resulted in a growing growth in the number of such databases [2, 3]. There is a great need to resort to intelligent approaches to get knowledge from processed data. As a result, data mining has become a popular research field [4, 5]. Classification is a supervised machine learning problem in which a collection of training data is used to map input data into one of several predetermined categories [5, 6]. Any classification algorithm's purpose is to create a technique

which can reliably predict the category of unobserved examples [7]. Classification has numerous uses in a range of areas, including document organization, medical diagnosis, and many more. Many classification techniques and models have been devised and used as a result of this, including radial basis function (RBF) network [8], naive Bayes (NB) classifier [9], support vector machines (SVMs) [10], and K-nearest neighbours (KNN) algorithm [11] and several others.

To solve classification difficulties, ANNs have been frequently used [12]. There are several different kinds of ANNs [13] as modular neural networks, RBF networks, feed-forward neural networks, learning vector quantization neural networks, and several others. Not only do the aforementioned ANNs differ in terms of how they apply to learning, but also in terms of their control method and topology [14].

The PNN is considered a feed-forward neural network that can be used to predict challenges and solve classification. The gradient steepest descent approach, a common optimization technique, is used in the PNN technique to minimize errors between the predicted and actual output functions by allowing the network to modify the weights of the network [15].

The goal of merging metaheuristic algorithms and NN to build classification tools like the PNN is to improve effectiveness and efficiency while also allowing for more accurate and faster solutions of complex problems.

The problem statement was specified by one basic research question: Can the searchability of the Material Generation Algorithm have the ability to choose the best weights so we can get the best accuracy?

Metaheuristics are split into two kinds: population-based and single-based. Genetic algorithm (GA) [16], particle swarm optimization (PSO) [17], water evaporation optimization (WEO) [18], differential evolution (DE) [19], firefly algorithm (FA) [20], artificial bee colony (ABC) [21] and several others are population-based metaheuristics. Local search (LS) [22], tabu search [23] and simulated annealing (SA) [24] are examples of single-based metaheuristics.

The Material Generation Algorithm (MGA) was investigated and utilized in this research to enhance the efficiency of the PNN in solving the classification problem [25]. The PNN was utilized to generate some preliminary solutions that were generated at random and the MGA was utilized to tune the weights of the PNN.

The study is divided as follows. Section II is showed a background and literature review for the MGA. In Section III, the background on big data and its issues is presented. The PNN approach is described in detail in Section IV, while the MGA is described in detail in Section V. The proposed methodology is then detailed in Section VI. In Section VII, the outcomes are presented. Finally, in Section VIII, the conclusion is offered.

II. BACKGROUND AND LITERATURE: MATERIAL GENERATION ALGORITHM (MGA)

The author in [25] suggested that MGA be used to solve engineering challenges in the best possible design. The MGA has identified some of the advanced and fundamental parts of materials chemistry as inspirational concepts, notably the formation of chemical molecules and chemical reactions in the production of new materials. This research demonstrates that the MGA is able to produce highly competitive, if not exceptional, outcomes that outperform other metaheuristics.

The author in [26] presented the optimal design of truss structures using the MGA. For statistical purposes, many optimization runs are carried out. The results showed that the MGA may produce extremely acceptable, resulting in the smallest potential weight compared with the outcomes of a number of metaheuristic methods.

The author in [27] optimized the moulding parameters of resin reinforced sand mould cores using a hybrid Taguchi-WASPAS- MGA to get the optimum outcomes.

The author in [28] used sunflower optimization algorithm and MGA for efficient generation and analysis of materials and equipment of mechanical reducer for the material handling industry. The results showed that the technique is precise in providing better output.

III. BIG DATA: OPPORTUNITIES AND CHALLENGES

The existence of trillions of records that have been produced by millions of people and kept in a variety of online sites suggests the concept of big data [28]. Scientists can use the big data to address issues with small data samples by giving adequate test data to evaluate models, better handling noisy train data, avoiding overfitting models to train data and loosening theoretical model assumptions. In Big Data, there are challenges, like capturing, transferring, storing, cleaning, analyzing, filtering, searching, sharing, securing, and visualizing data [29]. Different research communities have been battling to produce a dynamic, fast, new, and user-friendly Big Data technology [28], which contribute to solving many problems related to data and how to retrieve it.

IV. PROBABILISTIC NEURAL NETWORK (PNN)

The PNN was proposed for the first time in[30]. The training of a PNN does not entail using heuristic searches to

find the best smoothing factor, as this is an optimization problem [31]. A four-layered feed-forward network is formed: (a) input layer, (b) hidden layer, (c) summation layer, (d) output layer, using a statistical algorithm. Fig. 1 illustrates the architecture of a typical PNN. Each input neuron acts as a unique characteristic from the train and test datasets [32]. The PNN network's four levels are detailed below:

- Input layer: Each indicator variable is represented by a neuron. The categorical factors are made up of N-1 neurons, with N being the number of categories. By subtracting the middle value, the input neuron is expected to normalize the value range. It then divides it into quartile range values.
- Hidden layer: Each occurrence in the training dataset is represented by a single neuron. Each training sample has one unit which creates a product of the input vector x and the weight vector wi, zi = x.wi, and then runs the nonlinear procedure:

$$\exp \left[\left(\frac{(w_i - x)^T \cdot (-w_i - x)}{(2\alpha^2)} \right) \right] \quad (1)$$

- Pattern/summation layer: A single pattern neuron is available for each class of objective criteria. The weight value that emerges from the hidden neurons is given to the pattern neurons that match with the hidden neurons. Each training group's objective class is stored alongside each hidden neuron. which combines the contributions for each type of input and provides the output of a network as a probabilistic vector:

$$\Sigma_i \left[\left(\frac{(-w_i - x) \cdot (w_i - x)^T}{(2\alpha^2)} \right) \right] \quad (2)$$

- Output / Decision layer: creates binary classes that correspond to the decision classes Ω_s and Ω_r , $s \neq r$, $s, r = 1, 2, \dots, q$ relies on the following criteria of classification.

$$\Sigma_i \left[\left(\frac{(w_i - x)^T \cdot (-w_i - x)}{(2\alpha^2)} \right) \right] > \Sigma_j \left[\left(\frac{(w_j - x)^T \cdot (-w_j - x)}{(2\alpha^2)} \right) \right] \quad (3)$$

There is just one weight for these nodes, C, the number of training samples in each class and the prior membership probabilities, C, given by the cost parameter:

$$C = \frac{h_s l_s}{h_r l_r} \cdot \frac{n_r}{n_s} \quad (4)$$

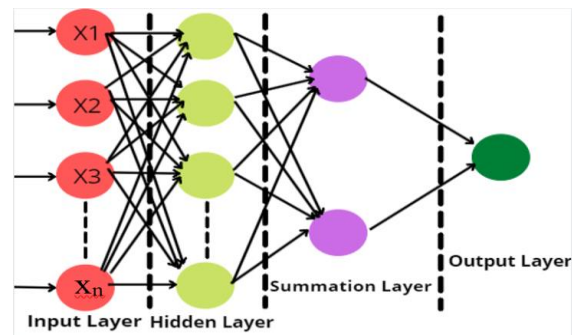


Fig. 1. Probabilistic Neural Network Structure.

V. MATERIAL GENERATION ALGORITHM (MGA)

In the year 2021, MGA is a bioinspired algorithm inspired by material chemistry[25]. To construct and formulate a well-defined mathematical model for the new method, the basic principles of chemical compounds, reactions, and stability are used. MGA determines a number of materials (Mat) made of several periodic table elements (PTEs), based on the fact that such natural evolution technique create a preset population of solution candidates that are evolved by random changes and selection. A materials numbers are examined as solution candidates (Mat_n) in this algorithm, each of which is made up of some constituents that are represented as decision variables (PTE_jⁱ). The following is the mathematical representation of these two components:

$$\text{Mat} = \begin{bmatrix} \text{Mat}_1 \\ \text{Mat}_2 \\ \vdots \\ \text{Mat}_i \\ \vdots \\ \text{Mat}_n \end{bmatrix} = \begin{bmatrix} \text{PTE}_1^1 & \text{PTE}_1^2 & \dots & \text{PTE}_1^j & \dots & \text{PTE}_1^d \\ \text{PTE}_2^1 & \text{PTE}_2^2 & \dots & \text{PTE}_2^j & \dots & \text{PTE}_2^d \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \text{PTE}_i^1 & \text{PTE}_i^2 & \dots & \text{PTE}_i^j & \dots & \text{PTE}_i^d \\ \vdots & \vdots & \dots & \vdots & \ddots & \vdots \\ \text{PTE}_n^1 & \text{PTE}_n^2 & \dots & \text{PTE}_n^j & \dots & \text{PTE}_n^d \end{bmatrix}, \quad \begin{cases} i=1,2,\dots,n. \\ j=1,2,\dots,d. \end{cases} \quad (5)$$

There are two variables in the mathematical equations where d denotes the number of items (decision variables) in each subject (the candidate solutions) and n denotes the total number of items considered.

PTE_jⁱ is determined at random in the first step of the optimization procedure, whereas the decision variable boundaries are defined based on the problem under consideration. The initial placements of PTEs in the search space are set at random:

$$\text{PTE}_j^i(0) = \text{Unif}(0,1) \cdot (\text{PTE}_{i,\max}^j - \text{PTE}_{i,\min}^j) + \text{PTE}_{i,\min}^j, \quad \begin{cases} i=1,2,\dots,n. \\ j=1,2,\dots,d. \end{cases} \quad (6)$$

Where PTE_jⁱ(0) is the beginning value of the jth element in the ith material; PTE_{i,min}^j and PTE_{i,max}^j are the minimum and maximum permissible values for the jth decision variable of the ith solution candidate, respectively; and Uni f(0, 1) is a random number in the [0, 1] range.

To mathematically simulate chemical compounds, all PTEs are considered to be in the ground state, which can be externally activated by magnetic areas, photon or light absorption, and interactions with other colliding entities or particles in the case of ions or other individual electrons. Elements have a tendency to gain, lose, or even share electrons with other PTEs due to their varied stabilities, resulting in ionic or covalent compounds. Using the initial Mat in equation (5), d random PTEs are chosen to model the ionic and covalent compounds. The probability theory is used to model the operations of sharing electrons, gaining, and losing for the selected PTEs. To achieve this goal, for each PTE, a continuous probability distribution is used to configure a

chemical molecule, which is then regarded a new PTE, as follows:

$$\text{PTE}_{\text{new}}^k = \text{PTE}_{r_1}^{r_2} \pm e^-, \quad k=1,2,\dots,d \quad (7)$$

R2 and r1 are random integers uniformly distributed in the intervals [1, d] and [1, n], respectively; PTE_{r1}^{r2} is from the Mat that was chosen at random; e- is the probabilistic component for simulating electron loss, gain, and sharing in the mathematical model represented with a normal Gaussian distribution; and PTE_{new}^k new is the new material. PTEs are used to construct a new material (Matnew1), which is being added as a new solution filter to the list of the raw material (Mat):

$$\text{Mat}_{\text{new1}} = [\text{PTE}_{\text{new}}^1 \text{ PTE}_{\text{new}}^2 \dots \text{PTE}_{\text{new}}^k \dots \text{PTE}_{\text{new}}^d], \quad k=1,2,\dots,d \quad (8)$$

The candidates for the overall solution are then integrated and displayed as follows:

$$\text{Mat} = \begin{bmatrix} \text{Mat}_1 \\ \text{Mat}_2 \\ \vdots \\ \text{Mat}_i \\ \vdots \\ \text{Mat}_n \\ \text{Mat}_{\text{new1}} \end{bmatrix} = \begin{bmatrix} \text{PTE}_1^1 & \text{PTE}_1^2 & \dots & \text{PTE}_1^j & \dots & \text{PTE}_1^d \\ \text{PTE}_2^1 & \text{PTE}_2^2 & \dots & \text{PTE}_2^j & \dots & \text{PTE}_2^d \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \text{PTE}_i^1 & \text{PTE}_i^2 & \dots & \text{PTE}_i^j & \dots & \text{PTE}_i^d \\ \vdots & \vdots & \dots & \vdots & \ddots & \vdots \\ \text{PTE}_n^1 & \text{PTE}_n^2 & \dots & \text{PTE}_n^j & \dots & \text{PTE}_n^d \\ \text{PTE}_{\text{new1}}^1 & \text{PTE}_{\text{new1}}^2 & \dots & \text{PTE}_{\text{new1}}^k & \dots & \text{PTE}_{\text{new1}}^d \end{bmatrix}, \quad \begin{cases} i=1,2,\dots,n. \\ j=1,2,\dots,d. \\ k=1,2,\dots,d. \end{cases} \quad (9)$$

Fig. 2 depicts the structure of the mentioned method for configuring new materials based on chemical components (covalent and ionic).

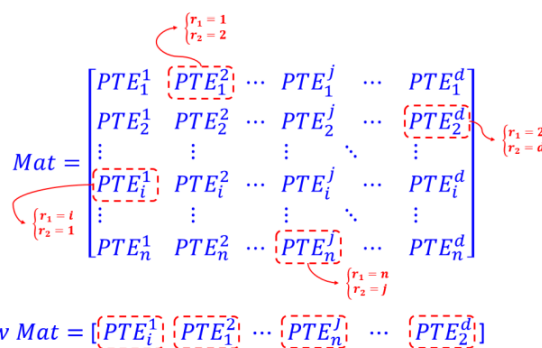


Fig. 2. The Random Periodic Table Elements (PTE) Selection and Creation of New Materials are Depicted Schematically.

The following is the probability of selecting a new element (PTE_{new}^k) in relation to the randomly picked first element (PTE_{r1}^{r2}):

$$f(\text{PTE}_{\text{new}}^k | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad k=1,2,\dots,d \quad (10)$$

The symbol for the standard deviation in the previous equation is σ; the symbol for the variance is σ²; μ is the

median, expectation of the distribution or mean, which corresponds to the randomly chosen PTE (PTE_{r1}^2); and e is the natural logarithm's Napierian base or natural base.

Chemical reactions are a type of manufacturing method in which various chemical changes are decided for producing products with altered characteristics that are even distinct from the initial reaction mixture. To simulate the procedure of manufacturing new materials mathematically using the reaction mixture idea, an integer random number (l) is determined depend on the materials in the first Mat that are examined for participation in a reaction mixture. After that, to decide the placements of the picked materials in the initial Mat , l integer random numbers (m_j) are created. As a result, new solutions are created that are linear combinations of the previous ones. Fig. 3 depicts a schematic illustration of the given procedure, with the following mathematical representation:

$$Mat_{new2} = \frac{\sum_{m=1}^l (p_m \cdot Mat_{m_j})}{\sum_{m=1}^l (p_m)}, j=1,2,\dots,l \quad (11)$$

The Mat_m is the m th randomly chosen material from the first Mat , Mat_{new2} is the new material created by the chemical reaction idea and p_m is the normal Gaussian distribution for the m th material participation factor.

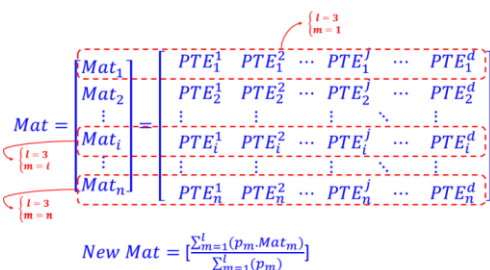


Fig. 3. A Diagram of the Random Material Selection Method for Developing New Materials.

VI. METHODOLOGY PROPOSED: MGA WITH PNN

The MGA was utilized in this research to determine the best weights to employ with the PNN algorithm. To address the classification issues, it suggested the MGA–PNN, a new hybrid method. As shown in Fig. 4, the method begins with the PNN producing the initial weights randomly. Following that, the input values are multiplied by the matching weights w_{ij} , which are based on the PNN model's values.

The proposed MGA–PNN structure is shown in Fig. 5. It is divided into two sections: the first is the PNN, which makes utilize the training data. The data that has been tested is then classified. The accuracy is calculated using equation (12). The MGA is then used to fine-tune the PNN weights. The new data will then be tested for accuracy. This method is continued until the end criteria have been fulfilled.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (12)$$

The object is categorized as TN if both the expected and actual labels are negative. The class is categorized as TP if both the expected and actual labels of the object are positive. Further, the class is categorized as FP, when the anticipated class is positive, but the actual label is negative. The

anticipated class is negative, but the actual label is positive, therefore it's categorized as FN. See Table I [33].

For evaluating the proposed MGA-PNN performance, three additional performance measurements were calculated: The rate of error was found (equation (13)), specificity (equation (14)), sensitivity (equation (15)) and G-mean (equation (16)).

$$Error\ Rate = 1 - \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

$$Specificity = \frac{TN}{FP+TN} \quad (14)$$

$$Sensitivity = \frac{TP}{FN+TP} \quad (15)$$

$$G\text{-mean} = \sqrt{(Sensitivity \times Specificity)} \quad (16)$$

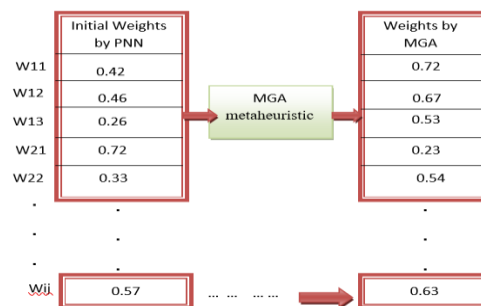


Fig. 4. Representation of Initial Weights.

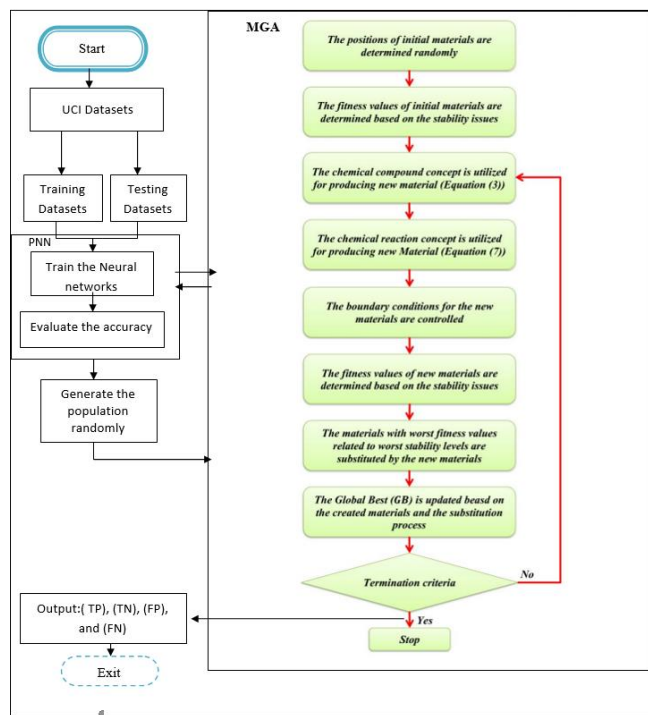


Fig. 5. Flowchart of MGA-PNN Technique.

TABLE I. CROSS-MATRIX CLASSIFICATION

	Positive	Negative
Positive	TP	FN
Negative	FP	TN

VII. EXPERIMENTS AND RESULTS

The efficiency of the MGA-PNN approach is measured in this research using 11 benchmark UCI datasets. We will compare the outcomes of proposed method (MGA-PNN) with PNN, biogeography-based optimization (BBO) and firefly algorithm (FA).

A. Description of the Dataset

These studies are based on a set of a datasets, which may be found at [7]. The previous link provides the size of the testing and training sets. The split was made using a basic train/test split algorithm with a training size is equal to 0.7 and a testing size is equal to 0.3.

B. The Categorization Quality Evaluation Results

Experiments are run on a Windows 10 professional PC with MATLAB R2015b and 16 GB RAM with an Intel ®

Xeon ®CPU ES-1630 v3 @3.70 GHz computer. Table II displays the settings for the input parameters.

The recommended method's rating quality is determined by their ability to improve the desired solution. Table III compares the performance of the proposed MGA technique with PNN, FA [34] and BBO in terms of ratio G-mean, error rate (%), specificity, accuracy and sensitivity.

TABLE II. INPUT PARAMETER SETTING

NCompan (Maximum number of initial Component)	100
Globalbest	0
Max_iteration	200
Population_size	50

TABLE III. CLASSIFICATION SPECIFICITY, ACCURACY, ERROR RATE, RATIO G-MEAN AND SENSITIVITY FOR PNN, FA, BBO AND MGA

Dataset	Technique	TP	FP	TN	FN	Accuracy	Sensitivity	Specificity	Error rate	Ratio g-mean
PID	PNN	35	28	90	39	65.104	47.30	76.27	34.89	60.06
	FA-PNN	33	30	113	16	76.040	67.35	79.02	24.00	72.95
	BBO-PNN	38	25	99	30	71.350	55.88	79.84	28.56	66.79
	MGA-PNN	164	14	43	28	82.80	85.05	75.0	17.2	78.87
HSS	PNN	44	12	6	15	64.93	74.58	33.33	35.07	49.86
	FA-PNN	54	2	10	11	83.12	83.08	83.33	16.88	83.20
	BBO-PNN	52	4	11	10	81.82	83.87	73.33	18.18	78.42
	MGA-PNN	71	3	94	5	95.38	93.42	96.91	4.62	95.15
AP	PNN	23	1	1	2	88.88	92.00	50.00	11.12	67.82
	FA-PNN	24	0	1	2	92.59	92.31	100.00	7.41	96.08
	BBO-PNN	52	4	11	10	81.82	83.87	73.33	18.18	78.42
	MGA-PNN	93	0	7	3	93.88	92.86	100.00	6.12	96.36
BC	PNN	14	9	36	13	69.44	51.9	80.00	30.60	64.44
	FA-PNN	31	1	24	12	80.88	72.09	96.00	19.12	83.19
	BBO-PNN	13	10	44	5	79.17	72.22	81.48	20.83	76.71
	MGA-PNN	17	6	44	5	84.72	77.27	88.00	12.28	82.46
LD	PNN	18	15	34	19	60.46	48.60	69.40	39.50	58.08
	FA-PNN	31	1	24	12	79.07	72.09	96.00	20.93	83.19
	BBO-PNN	32	0	23	13	72.09	71.11	100.00	27.91	84.30
	MGA-PNN	24	9	49	4	84.88	85.71	84.48	15.12	85.09
Heart	PNN	27	5	23	13	73.53	67.50	82.10	26.50	74.44
	FA-PNN	31	1	24	12	80.88	72.09	96.00	19.12	83.19
	BBO-PNN	32	0	23	13	80.90	71.11	100.00	19.10	84.33
	MGA-PNN	32	0	25	11	83.82	74.42	100.00	16.18	86.27
GCD	PNN	133	46	39	32	68.80	80.60	45.90	31.20	60.82
	FA-PNN	166	13	30	41	78.40	80.19	69.77	21.60	74.79
	BBO-PNN	139	40	44	27	73.20	83.73	52.38	26.80	66.23
	MGA-PNN	165	14	42	29	82.80	85.05	75.00	17.20	79.87
Parkinson's	PNN	39	0	4	6	87.75	86.67	100.00	12.25	93.09
	FA-PNN	38	1	6	4	89.80	90.48	85.71	10.2	88.06
	BBO-PNN	39	0	7	3	93.88	92.86	100.00	6.12	96.36
	MGA-PNN	93	0	7	3	93.88	92.86	100.00	6.12	96.36
SPECTF	PNN	49	4	5	9	80.59	84.48	55.56	19.40	68.51
	FA-PNN	52	1	10	4	92.54	92.86	90.91	7.46	91.88
	BBO-PNN	49	9	5	5	86.57	90.74	35.71	13.43	56.92
	MGA-PNN	49	4	12	2	91.04	96.08	75.00	8.96	84.88
ACA	PNN	60	14	84	15	83.24	80.00	85.70	16.8	82.80
	FA-PNN	65	9	94	5	91.91	92.86	91.26	8.09	92.06
	BBO-PNN	65	9	88	11	88.53	85.53	90.72	11.47	88.09
	MGA-PNN	24	9	49	4	84.88	85.71	84.48	15.12	85.09
Fourclass	PNN	59	19	127	11	86.11	84.29	86.99	13.89	85.63
	FA-PNN	78	0	138	0	100.00	100.00	100.00	0.00	100.00
	BBO-PNN	78	0	138	0	100.00	100.00	100.00	0.00	100.00
	MGA-PNN	78	0	138	0	100.00	100.00	100.00	0.00	100.00

The outcomes showed that the proposed algorithm is superior in 9 out of 11 datasets over the rest of the algorithms in Table III. The original PNN achieved 65.1% accuracy in the PIMA Indian diabetes (PID) dataset, while the proposed MGA-PNN attained 82.8 percent accuracy. All the best outcomes showed in bold. The suggested technique has strong exploitation capabilities and can come up with superior solutions because a large number of candidates are grouped around the best solution. On almost all datasets, the suggested MGA outperforms the original PNN approach in terms of error rate, sensitivity, specificity and accuracy.

The MGA's performance was further validated by examining whether it differed statistically from the FA. For classification accuracy, a t-test with a significance interval of 95 percent ($\alpha = 0.05$) was used. Table IV displays the suggested approach's standard deviations and accuracy means. The performance of the MGA is clearly superior to that of the FA, as all of the P-values are less than 0.01.

TABLE IV. THE P-VALUES FOR MGA ACCURACY WITH FA AND T-TEST ACCURACY

Dataset		Mean	Std.Deviation	Std.Error Mean	P-Value
PID	MGA	82.8000	0.00000	0.00000	0.00
	FA	73.4895	1.28560	0.23472	
HSS	MGA	94.5087	0.75519	0.13788	0.00
	FA	81.8179	1.02322	0.18681	
AP	MGA	92.5170	1.72282	0.31450	0.00
	FA	92.5926	0.00012	0.00002	
BC	MGA	82.9167	1.59613	0.29141	0.00
	FA	77.3935	1.74347	0.31831	
LD	MGA	83.3333	2.23006	0.40720	0.00
	FA	75.5810	1.49604	0.27310	
Heart	MGA	82.2549	1.80867	0.33022	0.00
	FA	78.6819	2.23781	0.40857	
GCD	MGA	82.8000	0.00000	0.00000	0.00
	FA	75.1600	1.58040	0.28854	
Parkinson's	MGA	92.5170	1.72282	0.31450	0.00
	FA	89.7950	0.00000	0.00000	
SPECTF	MGA	88.2668	1.37125	0.25035	0.00
	FA	88.8057	1.82787	0.33372	
ACA	MGA	83.3333	2.23006	0.40720	0.00
	FA	89.8840	1.05983	0.19350	
Fourclass	MGA	100.000	0.00000	0.00000	0.00
	FA	100.000	0.00000	0.00000	

VIII. CONCLUSION

The major goal of this study was to propose a new strategy for determining high-quality answers to categorization problems. The Material Generation Algorithm is a population-based metaheuristic that MGA is a bioinspired algorithm inspired by material chemistry. Therefore, the weight values of

the PNN can be optimized by MGA. When a huge search space is being examined, the MGA's superior exploitation and exploration capabilities allow it to achieve better results than FA and BBO. The MGA was utilized to tune the weight of the PNN in this study. To attain the research's targets, the results of this strategy rely on PNN and MGA was used to compare with the results original PNN's classification accuracy, FA-PNN and BBO-PNN. The MGA, which optimized the PNN weights, was used to improve the initial solutions, which were created randomly using the PNN. According to experimental results utilizing 11 benchmark datasets, the suggested MGA with PNN outperformed the original PNN, FA-PNN and BBO-PNN on 9 out of 11 benchmark datasets. This leads us to the fact that MGA can be implemented in additional real and high dimensional datasets to investigate their behavior under different situations in terms of trait numbers. As a result, we'll be focusing our efforts on this topic in the future.

IX. DISCUSSION

This study is considered one of the most important studies in the world of Data Mining. As our use of the method of merging with the metaphysical algorithms, especially with the MGA, and comparing its results with the results of 3 other studies (PNN, FA and BBO) that gives clear evidence of its importance in terms of increasing classification accuracy. As this study only used one algorithm to combine it with PNN, I believe that merging more than one of the high-specification meta-historical algorithms with PNN leads to an improvement and a significant increase in accuracy, and this is our destination in the work of these studies in the future.

REFERENCES

- [1] M. O. Wedyan, "Augmented reality and novel virtual sample generation algorithm based autism diagnosis system," 2020.
- [2] K. Khatatneh, I. Khataleen, R. Alshwaiyat, and M. Wedyan, "A Novel Student Clustering Model for the Learning Simplification in Educational Environments," vol. 10, no. 8, 2019.
- [3] J. Atwan et al., "The effect of using light stemming for Arabic text classification," vol. 12, no. 5, 2021.
- [4] M.-S. Chen, J. Han, and P. S. Yu, "Data mining: an overview from a database perspective," vol. 8, no. 6, pp. 866-883, 1996.
- [5] A. Khtoom and M. Wedyan, "Feature Selection Models for Data Classification: Wrapper Model vs Filter Model," in International Conference on Information, Communication and Computing Technology, 2019, pp. 247-257: Springer.
- [6] M. Wedyan, A. Al-Jumaily, and A. Crippa, "Using machine learning to perform early diagnosis of autism spectrum disorder based on simple upper limb movements," vol. 15, no. 4, pp. 195-206, 2019.
- [7] M. Wedyan, H. Al-Zoubi, and J. Atwan, "Error Detection and Correction Using a Genetic Algorithm," vol. 8, no. 2, pp. 9-16, 2019.
- [8] R. J. Howlett and L. C. Jain, Radial basis function networks 2: new advances in design. Physica, 2013.
- [9] S. Xu, "Bayesian Naïve Bayes classifiers to text classification," vol. 44, no. 1, pp. 48-59, 2018.
- [10] I. Steinwart and A. Christmann, Support vector machines. Springer Science & Business Media, 2008.
- [11] O. Sutton, "Introduction to k nearest neighbour classification and condensed nearest neighbour data reduction," vol. 1, 2012.
- [12] M. Wedyan, A. Al-Jumaily, and A. Crippa, "Early diagnose of autism spectrum disorder using machine learning based on simple upper limb movements," in International conference on hybrid intelligent systems, 2018, pp. 491-500: Springer.

- [13] M. Wedyan, A. Crippa, and A. Al-Jumaily, "A novel virtual sample generation method to overcome the small sample size problem in computer aided medical diagnosing," vol. 12, no. 8, p. 160, 2019.
- [14] C. Gershenson, "Artificial neural networks for beginners," 2003.
- [15] G. P. Zhang and P. C. Cybernetics, "Neural networks for classification: a survey," vol. 30, no. 4, pp. 451-462, 2000.
- [16] G. R. Harik, F. G. Lobo, and D. E. Goldberg, "The compact genetic algorithm," vol. 3, no. 4, pp. 287-297, 1999.
- [17] F. Marini, B. J. C. Walczak, and I. L. Systems, "Particle swarm optimization (PSO). A tutorial," vol. 149, pp. 153-165, 2015.
- [18] A. Kaveh, T. Bakhshpoori, and Structures, "Water evaporation optimization: a novel physically inspired optimization algorithm," vol. 167, pp. 69-85, 2016.
- [19] K. V. Price, "Differential evolution," in Handbook of optimization: Springer, 2013, pp. 187-214.
- [20] X.-S. Yang, "Firefly algorithms for multimodal optimization," in International symposium on stochastic algorithms, 2009, pp. 169-178: Springer.
- [21] F. S. Abu-Mouti and M. E. El-Hawary, "Overview of Artificial Bee Colony (ABC) algorithm and its applications," in 2012 IEEE International Systems Conference SysCon 2012, 2012, pp. 1-6: IEEE.
- [22] V. Kalogeraki, D. Gunopulos, and D. Zeinalipour-Yazti, "A local search mechanism for peer-to-peer networks," in Proceedings of the eleventh international conference on Information and knowledge management, 2002, pp. 300-307.
- [23] F. Glover and M. Laguna, "Tabu search," in Handbook of combinatorial optimization: Springer, 1998, pp. 2093-2229.
- [24] P. J. Van Laarhoven and E. H. Aarts, "Simulated annealing," in Simulated annealing: Theory and applications: Springer, 1987, pp. 7-15.
- [25] S. Talatahari, M. Azizi, and A. H. Gandomi, "Material generation algorithm: a novel metaheuristic algorithm for optimization of engineering problems," vol. 9, no. 5, p. 859, 2021.
- [26] M. Azizi, M. B. Shishehgarkhaneh, and M. Basiri, "Optimum design of truss structures by Material Generation Algorithm with discrete variables," vol. 3, p. 100043, 2022.
- [27] N. C. Behera, S. Jeet, C. K. Nayak, D. K. Bagal, S. N. Panda, and A. Barua, "Parametric appraisal of strength & hardness of resin compacted sand castings using hybrid Taguchi-WASPAS-Material Generation Algorithm," vol. 50, pp. 1226-1233, 2022.
- [28] S. Jena, S. Jeet, D. K. Bagal, A. K. Baliarsingh, D. R. Nayak, and A. Barua, "Efficiency analysis of mechanical reducer equipment of material handling industry using Sunflower Optimization Algorithm and Material Generation Algorithm," vol. 50, pp. 1113-1122, 2022.
- [29] R. Bedeley and L. Iyer, "Big Data opportunities and challenges: the case of banking industry," in Proceedings of the Southern Association for Information Systems Conference, 2014, vol. 1, pp. 1-6.
- [30] D. F. Specht, "Probabilistic neural networks," vol. 3, no. 1, pp. 109-118, 1990.
- [31] M. Wedyan, O. Elshaweesh, E. Ramadan, and R. Alturki, "Vibrating Particles System Algorithm for Solving Classification Problems," vol. 43, no. 3, pp. 1189-1206, 2022.
- [32] W. Sweeney, M. Musavi, and J. Guidi, "Probabilistic neural network as chromosome classifier," in Proceedings of 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan), 1993, vol. 1, pp. 935-938: IEEE.
- [33] M. Alweshah, E. Ramadan, M. H. Ryalat, M. Almi'ani, and A. Hammouri, "Water evaporation algorithm with probabilistic neural network for solving classification problems," vol. 6, no. 1, pp. 1-15, 2020.
- [34] M. Alweshah and S. Abdullah, "Hybridizing firefly algorithms with a probabilistic neural network for solving classification problems," vol. 35, pp. 513-524, 2015.