*Computer Systems Science & Engineering*
DOI:10.32604/csse.2022.024210
***Article***

Tech Science Press

# Vibrating Particles System Algorithm for Solving Classification Problems

**Mohammad Wedyan[1], Omar Elshaweesh[2], Enas Ramadan[3] and Ryan Alturki[4,*]**

[1]Department of Autonomous Systems, Faculty of Artificial Intelligence, Al-Balqa Applied University, Al-Salt, 19117, Jordan
[2]Department of Software Engineering, Information Technology College, AL-Hussein bin Talal University, Ma'an, 71111, Jordan
[3]Department of Computer Science, Prince Abdullah Bin Ghazi Faculty of Communication and Information Technology, Al-Balqa Applied University, Al-Salt, 19117, Jordan
[4]Department of Information Science, College of Computer and Information Systems, Umm Al-Qura University, Makkah, 21961, Saudi Arabia
*Corresponding Author: Ryan Alturki. Email: rmturki@uqu.edu.sa
Received: 09 October 2021; Accepted: 09 December 2021

**Abstract:** Big data is a term that refers to a set of data that, due to its largeness or complexity, cannot be stored or processed with one of the usual tools or applications for data management, and it has become a prominent word in recent years for the massive development of technology. Almost immediately thereafter, the term "big data mining" emerged, i.e., mining from big data even as an emerging and interconnected field of research. Classification is an important stage in data mining since it helps people make better decisions in a variety of situations, including scientific endeavors, biomedical research, and industrial applications. The probabilistic neural network (PNN) is a commonly used and successful method for handling classification and pattern recognition issues. In this study, the authors proposed to combine the probabilistic neural network (PPN), which is one of the data mining techniques, with the vibrating particles system (VPS), which is one of the metaheuristic algorithms named "VPS-PNN", to solve classification problems more effectively. The data set is eleven common benchmark medical datasets from the machine-learning library, the suggested method was tested. The suggested VPS-PNN mechanism outperforms the PNN, biogeography-based optimization, enhanced-water cycle algorithm (E-WCA) and the firefly algorithm (FA) in terms of convergence speed and classification accuracy.

**Keywords:** Vibrating particles system (VPS); probabilistic neural network (PNN); classification problem; data mining

## 1 Introduction

Information technology's widespread use and accessibility led to the preemptive inflation of the volume of information in a way that had not been witnessed before, which has sparked debate about big data on the internet, in terms of the feasibility of its existence in this random image.

When we talk about big data, we're referring to massive amounts of data from a variety of sources that can be hundreds of terabytes or even petabytes in size [1]. Receiving email, browsing websites, doing online transactions, using Google Docs, uploading pictures, and other activities generate a large amount of data on a daily basis. Smart living, big data and working environments have been increasingly popular in recent years [2].

From here, the so-called data mining emerged in the late eighties as a method aimed to extract knowledge from massive amounts of data using mathematical algorithms, which form the foundation of data mining and can be found in a variety of disciplines such as artificial intelligence, logic, mathematics, learning science, and pattern recognition science, statistics and machine science by transforming it from a collection of incomprehensible facts into useful knowledge that can be exploited and benefited from later [1,3,4].

Data mining is a term used for the extraction of information from large amounts of data. It is utilized in many domains such as corporate analysis, fraud detection and market analysis due to its superior capability to extract knowledge. Data mining applications have begun to grow significantly because (1) the amount of available data is growing exponentially and (2) there is intense competition in the market that is pushing companies to make the most of their data. The purpose of data mining is to make a model that can be applied in classification, prediction or any other similar tasks [5].

Data mining is utilized in prediction, where some variables are utilized to predict other variables (classification), and in description, where patterns are specified that are easy to understand by the user (clustering) [6].

The classification operation divides data into independent categories; The goal is to get a precise prediction of the target group. Classification of data supports in producing the demanded output that can be utilized in the future. It is very significant because It's easy to decide on a lot of areas such as marketing, science and business [7–9].

The first step for a classification researcher is to choose an effective classification method. Researchers are currently concentrating their efforts on using the NN to address a variety of categorization challenges. The NN is commonly regarded as the most well-known and widely used categorization system, and its popularity originates from the fact that it closely resembles the human brain in terms of how knowledge and information are stored and acquired. This knowledge and information is used to recognize distinct patterns and forms in order to deliver appropriate reactions to specific activities [10,11].

The Probabilistic neural network (PNN) is a general artificial neural network (ANN) procedure that is dependent on the 'gradient steepest descent approach,' which allows the network to precisely change the network weights to reduce any errors between the estimated and real output jobs.

The purpose of merging metaheuristic algorithms and NNs to build distribution tools like the PNN is to improve efficiency and effectiveness while also allowing for the faster and more accurate solution of complex problems [12].

Metaheuristics are divided into two types: single-based and population-based. Tabu search (TS) [13], simulated annealing (SA) [14], and local search [15] are examples of single-based metaheuristics (LS). The firefly algorithm (FA) [16], genetic algorithm (GA) [17], particle swarm optimization (PSO) [18], artificial bee colony (ABC) [19], and many others are population-based metaheuristics.

In addition, population-based techniques to optimized NN weight problems have been developed by researchers. The population-based approach's core notion is that the algorithms iteratively enhance a set of solutions. These methods, on the other hand, have certain drawbacks, such as a focus on exploration rather than exploitation and a slow convergence speed.

Later, there are some of the most notable works related to metaheuristic with optimization and classification problem domain. [20] proposed a novel method in which a model based on optimized deep neural networks is utilized to forecast the battery life of IoT gadgets. [21] proposed a method for implementing the Gated Recurrent Unit (GRU) and Long-Short Term Memory (LSTM) approaches and compared the results to a genetically trained neural network. Its goal is to improve the training of recurrent neural networks so that they can make better predictions. [22] suggested a model depend on rough sets for lowering attributes and a fuzzy logic system for classification is proposed for the prediction of Heart and Diabetes illnesses. The experiment is carried out on a variety of heart disease datasets from the UCI Machine Learning library, including Hungarian, Cleveland, and Switzerland datasets, as well as a diabetic dataset from an Indian hospital. The results of the experiments reveal that the suggested prediction algorithm surpasses existing methods in terms of accuracy, sensitivity and specificity. [23] focuses on battery life estimates in IoT frameworks in the marine environment. The data is then loaded into a Deep Neural Network (DNN) model for optimum prediction outcomes, and rough set theory is applied for feature extraction. The results showed its superiority in terms the performance.

Metaheuristic algorithms have been used with several types of classifiers, resulting in improved performance when compared to traditional classification methods. These mechanisms have some limitations. For example, because they are more interested with exploration than with exploitation, they may have a slow convergence speed.

The Vibrating particles system (VPS) has become an important tool for tackling issues in a variety of fields, including robotics, engineering, business optimization and image processing. Furthermore, the VPS's simplicity, flexibility, and adaptability make it effective at solving a wide range of real-world issues.

In this study, we are hybridized the VPS with PNN to tune the PNN weights to allow the system to deliver the best results possible using the weights collected from the PNN and increase the classification system's performance. To evaluate the proposed method's performance, it is checked on eleven benchmark distribution issues from the machine-learning repository (UCI).

## 2 Big Data

Big data plays a significant role in our lives and communities in the future. Governments have resorted to data mining within the contents of transactions and social networks represented by the Internet, identifying information related to the country's security, and detecting suspicious groups. Also, in our time, there are a lot of electronic service providers in the sale of products, and service providers track customer behaviors via the Internet and monitor customer satisfaction to increase profits and improve marketing efforts [24].

The gap between the linkages of capabilities that existing database management systems and big data management have reached an all-time high. Big data's three Vs (velocity, variety, and volume) each represent a different facet of today's DBMSs' critical flaws; Huge volumes necessitate tremendous scalability and parallelism, both of which are beyond the capabilities of today's DBMSs; Big data's wide range of data kinds defies the limitations of today's database systems' restricted processing architecture [25]. The speed/velocity demand of large data processing necessitates real-time efficiency that is well above the capabilities of current DBMSs. Current DBMSs' limited availability defeats big data's velocity request from a different perspective (Most current DBMSs require that data be imported/loaded into their storage systems in a standard format before any access/processing is permitted. When dealing with vast amounts of massive data, the importing/loading stage could take hours, days, or even months. As a result, the DBMSs' availability is significantly delayed/reduced).

Several attempts at using enormous parallel processing architectures to confront the scalability difficulty of big data have been made. Google was the first to make such an endeavor. Google produced the MapReduce programming methodology [26], which was combined with (and aided by) the GFS (Google File System [27]), a distributed file system that allows data to be readily partitioned among thousands of nodes in a cluster. Hadoop MapReduce is an Apache open-source version of Google's MapReduce architecture produced by Yahoo and other large corporations. Users can write two functions, reduce and map, in order to handle a large number of data entries occurring at the same time using the MapReduce framework [28].

## 3 Data Mining

Data mining is a term used for the extraction of information from huge amounts of data. It is utilized in many domains such as corporate analysis, fraud detection and market analysis due to its superior capability to extract knowledge. The applications of data mining have begun to grow significantly because (1) the amount of available data is growing exponentially and (2) there is intense competition in the market that is pushing companies to make the most of their data [29].

In most cases, through data mining, important patterns and relationships are obtained within a big amount of raw data and through these results can be applied to generate future predictions in the world. Data mining is used in a multitude of areas such as research, medicine, business and engineering [30].

The application of techniques and algorithms for data mining has faced many challenges due to the insufficient scalability of these techniques and algorithms. The volume of data that is created is very large, as knowledge is often obtained on an ongoing basis that needs mining and processing in (nearly) real time [30,31]. The utility of even extremely important knowledge is invalidated when it is discovered late. Big data presents new difficulties as well as opportunities, as it connects disparate data sets with heterogeneous and complex content, revealing new sources of information and insights. If we don't have the necessary mechanisms to control big data's "wildness", it will mutate a monster that is useless. We believe that big data should be viewed as a vastly increased resource for humans. All that's left is to provide the necessary tools for storing, accessing, and analyzing data efficiently (SA2 for short). Data mining algorithms and approaches are used to overcome some obstacles, but it not able to all challenges [32]. Big data mining necessitates enhanced parallel computing environments, improved preprocessing procedures such data filtering and integration, highly scalable algorithms and techniques and smart user interaction [33].

## 4 Big Data Mining

Big data mining approaches aim for more than just retrieving desired data or revealing hidden links and patterns between numerical parameters. Analyzing big amounts of data in a short amount of time can lead to new discoveries and theoretical conceptions [34]. When compared to the results obtained from mining traditional datasets, the disclosure of a big amount of unconnected big data leads to enhances our knowledge in the target area [35].

However, the research community faces a number of additional obstacles as a result of this. If challenges are overcome, this leads to the production of a combination of leading-edge data, algorithms and technologies. Use massively parallel computer architectures is one feasible option to improve existing approaches and algorithms [30].

## 5 Data Mining Techniques

Data mining is a vast discipline with numerous applications, therefore it has become an attractive subject to research [32]. The three categories of data mining approaches are depiction, hypothesis, and association. Data mining techniques can be utilized in a set of ways, and while their use is controversial, they can be beneficial if used correctly. A few data mining approaches are described and briefly discussed below:

### *5.1 Classification*

The most crucial phase in classification is to choose the right mechanism for the job. There are a variety of methods for resolving classification issues. The classification of data aids in the production of a desired output that may be reused in the future. It's crucial because it helps people make better decisions in fields like science, marketing, biomedicine, and business [36]. Many of the categorization techniques used in data mining are artificial intelligence-based. There are a number of techniques: the support vector machine (SVM) [37], logistic regression (LR) [38], radial basis function (RBF) [39], and neural network (NN) [40].

### *5.2 Clustering*

Data can be stored as huge data in a physical or computational structure [41]. Different vaults are utilized to store such data. The term "big data" refers to a data set that exceeds the computing capabilities of software. Clustering is the process of assembling multiple goods and their classifications depending on several aspects such as area, association, and so on. Schools, for example, might be grouped together based on similarities or differences. Understudies can also be classified according to their behavior. Clustering is used to search through data that is routinely collected [42].

### *5.3 Prediction*

Predictions are frequently based on prior experience and knowledge. Predictor variable refers to the emphasis on a specific piece of data as opposed to another piece of data. Prediction is a technique for predicting an ambiguous event based on prior knowledge or experience [43].

### *5.4 Text Mining*

Text data is explicit using this data mining method, which is depicted as text data in data mining. Records, communications, and html documents are examples of text data. Record processing, report summary, indexing, subject clustering, and mapping are some of the tasks that can be delegated to text mining. It is most commonly used in business and training. Associations have a large number of records and utilize text mining to get the data they need. Publishing, media distribution, information technology (IT), banking, government agencies, and pharmaceutical businesses are all examples of text mining applications [43].

## 6 Challenges of Data Mining

There are a number of issues that can be viewed as bottlenecks in the development of data mining techniques. A few of the most important challenges are listed here.

- ***Assessing statistical significance and overfitting:***

Data sets for data mining are compiled from a variety of sources, resulting in misleading data sets. As a result, a variety of regularization approaches and sampling techniques are utilized to plan data mining models [2,43].

- ***High-dimensional data set from Larye:***

The algorithms of data mining must cope with massive a large volume of data, both in terms of dimension and size. As a result, faster and more productive algorithms are necessary to handle this massive a large volume of data. As a result, in order to process this massive a large volume of data, faster and more productive algorithms are required. Parallel processing, partitioning, sampling, and other techniques are examples of possible configurations [43].

- ***Prior knowledge and interaction with the user:***

Data mining is an approach that requires interaction and iteration. Users must interact with the system at various stages. Domain knowledge can be applied at both a high level and a more detailed level when determining the model [43].

## 7 Background

### *7.1 Probabilistic Neural Network (PNN)*

The PNN is one of the classification techniques that consider most efficient and effective. Over other methods, the PNN provides the following benefits: (i) the training is completed quickly; (ii) it has a low sensitivity to outliers; (iii) it outperforms the MLP-PNN in terms of accuracy; and (iv) it is capable of generating precise target probability scores [44]. However, there are two drawbacks: (i) a significant memory requirement and (ii) slow execution [6]. As illustrated in Fig. 1, the PNN is organized as a four-layer feed-forward network structure with an input layer, summation layer, pattern layer, and output layer [45,46]. The PNN network's four levels are detailed below:

- ***The input layer:*** which feeds values to each of the pattern layer's neurons through a predictive variable for each neuron [47].
- ***The pattern layer:*** each training sample has one unit which creates a product of the input vector x and the weight vector wi, $zi = x.w^t_i$, and then runs the nonlinear procedure [47]:

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

- ***A summation layer:*** which combines the contributions for each type of input and provides the output of a network as a probabilistic vector [47]:

$$\sum_i\left[\left(\frac{(-\mathrm{w_I}-\mathrm{x}).(\mathrm{w_I}-x)^{\mathrm{T}}}{(2\alpha^2)}\right)\right] \tag{2}$$

- ***An output layer:*** creates binary classes that correspond to the decision classes $\Omega_r$ and $\Omega_s$, $r \neq s$, r, s = 1, 2, …, q relies on the following criteria of classification [47]:

$$\sum_i\left[\left(\frac{(-\mathrm{w_i}-\mathrm{x}).(\mathrm{w_i}-x)^{\mathrm{T}}}{(2\alpha^2)}\right)\right] > \sum_j\left[\left(\frac{(-\mathrm{w_j}-\mathrm{x}).(\mathrm{w_j}-x)^{\mathrm{T}}}{(2\alpha^2)}\right)\right] \tag{3}$$

These nodes only have one weight, C, which is determined by the prior membership probability and the training samples number in every class, which is determined by the cost parameter [48]:

$$C = -\frac{\mathrm{h_s l_s}}{\mathrm{h_r l_r}}.\frac{n_r}{n_s} \tag{4}$$

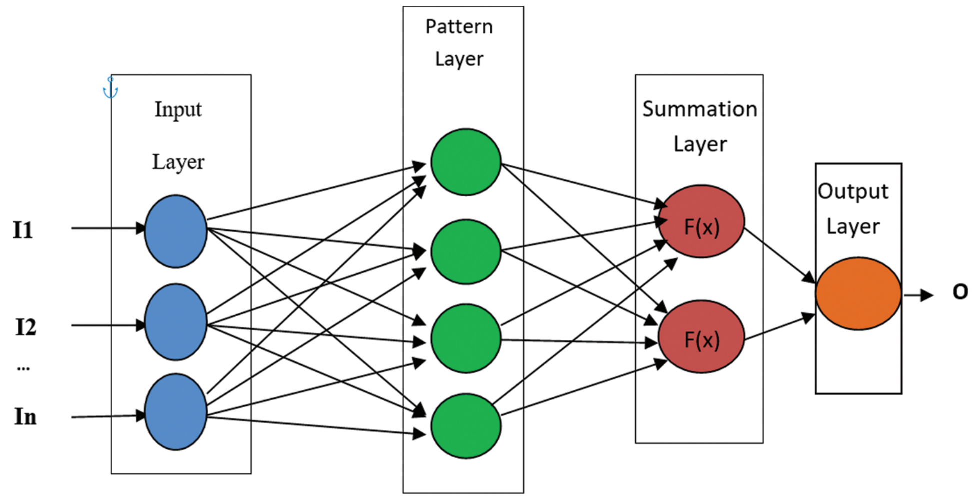


**Figure 1:** Probabilistic neural network structure

### *7.2 Vibrating Particles System (VPS)*

The VPS is a modern metaheuristic and population-based technique. In free vibration, it mimics single-degree-of-freedom systems with viscous damping. The particles gradually near to their equilibrium positions that are finished from historically best position and current population to achieve an appropriate balance of exploitation and exploration. The level of particle quality is improved iteratively like the optimization procedure by using a combination of randomization and exploitation of the given outcomes [49].

The starting agents in this meta-heuristic process are formed in an acceptable range by [49]:

$$x_i^j = x_{min} + rand.(x_{max} - x_{min}) \tag{5}$$

e is the jth factor of the ith The particles make up the search space allowed from the starting points to the ends for the jth variable and rand is a random number in the range of [0, 1].

Three parameters are defined in the VPS algorithm (HB, GB and BP) [49]:

1. Until that iteration, HB (the historically best position of the total population) is the greatest candidate.
2. In each cycle, GB (a good particle) is chosen at random from among the partially best replies.
3. Each iteration selects BP (a bad particle) at random from among the partially worst replies.

Eq. (6) is used to define a descending function. This parameter is included in the vibration model to account for the effect of the damping level [49].

$$D = \left(\frac{iter}{iter_{max}}\right)^{-\alpha} \tag{6}$$

where iter is the current iteration's number, is the number of iterations in total, and is a constant. The following equation is used to create the next agents in the VPS algorithm [49]:

$$x_i^j = r_1.[D.A.rand1 + HB^j] + r_2.[D.A.rand2 + GP^j] + r_3.[D.A.rand3 + BP^j] \tag{7}$$

$$S = [r_1.(HB^j - x_i^j)] + [r_2.(GP^j - x_i^j)] + [r_3.(BP^j - x_i^j)] \tag{8}$$

$$r_1 + r_2 + r_3 = 1 \tag{9}$$

where $r_1$, $r_2$, and $r_3$ represent the relative importance of HB, GB, and BP, and rand1, rand2, and rand3 are uniformly distributed random values in the range [0, 1]. To speed up the convergence of the VPS method, a parameter called p (between 0 to 1) is established. When this argument is compared to rand, and if rand > p, then $r_3 = 0$ and $r_2 = 1 - r_1$. Fig. 2 shows a pseudocode for the vibrating particles system algorithm, which was adapted from Kaveh et al. [50].

```
Initialize the parameters of the mechanism
The starting locations are generated at random.
The objective function values are evaluated and HB is stored
While the max-number of iterations is not reached
        for-each particle
            The BP and GP are picked
            if rand > p
            r3=0 and r2=1-r1
            end-if
            for-each element
            New-location is gained by equation (5)
            end-for
           Harmony search mechanism is used to replenish harmed elements.
      end-for
      The objective function's values are assessed, and HB is modified.
      end-while
end-mechanism
```

**Figure 2:** The vibrating particles system's pseudocode

### *7.3 Methodology Proposed: VPS with PNN*

The PNN is hybridized with the VPS for the first time in this study with the aim of developing a more efficient solution to tackle classification problems, adjusting the PNN's weights and improving the accuracy of classification by better exploring and exploiting the search region and thereby solving a variety of classification issues in an effective way. The combination of the VPS and the PNN will be referred to as VPS-PNN from now on.

The PNN classifier generates the initial weights at random, as shown in Fig. 3. The values of the input data are multiplied by the associated weights, as calculated by the PNN classifier (ij). The approach begins with starting weights created at random by the original PNN classification mechanism, as shown in Fig. 4. The input data's values also are multiplied by the appropriate weights $w(ij)$, are decided by the PNN mechanism (see Fig. 5) and communicated to the pattern layer according to Eq. (3). The latter are transformed into summation and output layers using a transfer function, as shown in Eq. (4). Because only one output is commonly requested, the output layer normally only has one class. The purpose of the training phase is to figure out the most precise weights to assign to the connecting line. Additionally, the result is calculated periodically throughout training and the outcome from the training and testing datasets is matched to the preferred outcome. Fig. 3 depicts the flowchart for training a NN.

The processes for using the VPS with the PNN are depicted in Fig. 5. It is divided into two sections: The PNN is the initial part, which utilized the training data to classify the data that is being tested. Then, The PNN weights are tuned using the VPS. After that, the classified data procedure's accuracy is checked again until the criteria for termination is reached. The suggested technique's classification quality is assessed by computing the accuracy value by Eq. (10), which is relies on the count of false-negative (FN), false-positive (FP), true-negative (TN), and true-positive (TP) outcomes.

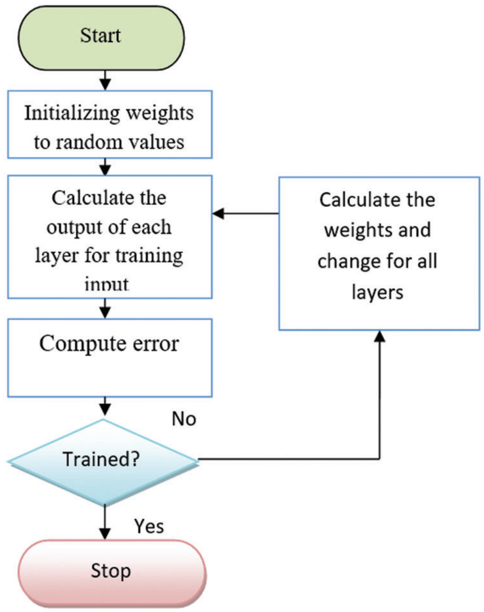


**Figure 3:** Flowchart of training NN

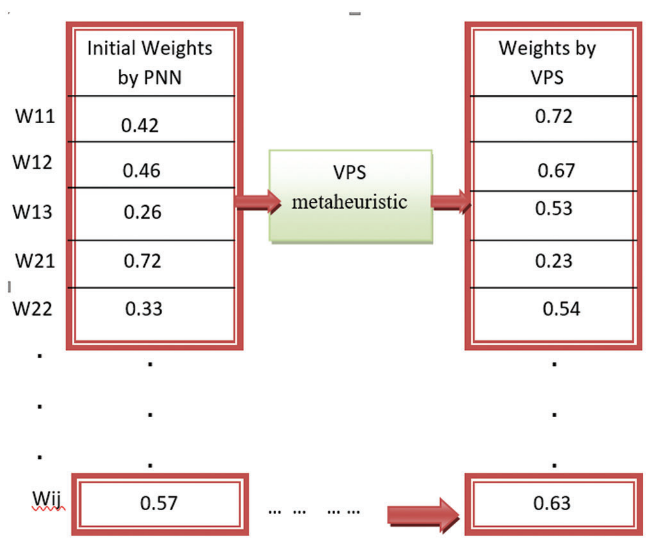


**Figure 4:** Initial weights

$$Accuracy = \frac{\text{True\_Negative} + \text{True\_Positive}}{\text{False\_Negative} + \text{False\_Positive} + \text{True\_Positive} + \text{True\_Negative}} \quad (10)$$

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 Tab. 1 below.

To evaluate the proposed VPS-PNN performance, three additional performance measurements were calculated: The rate of error was found (Eq. (11)), specificity (Eq. (12)), sensitivity (Eq. (13)) and G-mean (Eq. (14)).

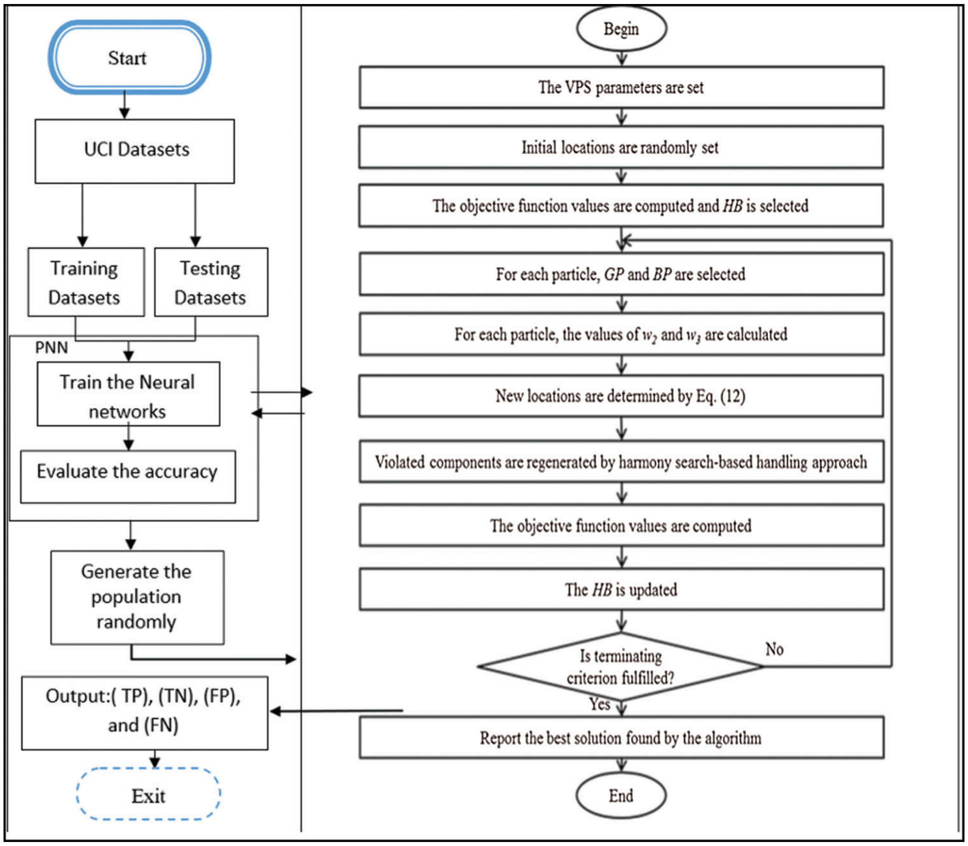


**Figure 5:** VPS-PNN mechanism flowchart

**Table 1:** Cross-matrix classification

| Actual category | Positive | Negative |
|---|---|---|
| Positive | True_positive | False_negative |
| Negative | False_positive | True_negative |

$$\text{Error Rate} = 1 - \frac{\#\text{of\_True\_Positive} + \#\text{of\_True\_Negative}}{\#\text{of\_True\_Positive} + \#\text{of\_True\_Negative} + \#\text{of\_False\_Positive} + \#\text{of\_False\_Negative}} \quad (11)$$

$$\text{Specificity} = \frac{\#\text{of\_True\_Negative}}{\#\text{of\_True\_Negative} + \#\text{of\_False\_Positive}} \quad (12)$$

$$\text{Sensitivity} = \frac{\#\text{of\_True\_Positive}}{\#\text{of\_True\_Positive} + \#\text{of\_False\_Negative}} \quad (13)$$

$$\text{G} - \text{mean} = \sqrt{(\text{Sensitivity} \times \ \text{Specificity})} \quad (14)$$

## 8 Experiments and Results

Experiments of the suggested approach were accomplished by using Matlab R2010a on a Windows 10 operating system and an Intel ® Xeon ®CPU ES-1630 v3 @3.70 GHz computer with 16 GB RAM. Tab. 2 lists the input parameters were utilized in all of the datasets and experiments.

**Table 2:** Input parameter setting

| Parameter | Value |
|---|---|
| Alpha | 0.05 |
| nVar (#of optimization variables) | 20 |
| xMin | −500 |
| xMax | 500 |
| #of_iterations | 200 |
| Size of the population | 20 |

The solutions were supplied in terms of best accuracy after 30 autonomous runs for each of the 11 datasets that may be freely obtained from http://csc.lsu.edu/huypham/HBA CBA/datasets.html. When the error is zero and the accuracy is 100%, it signifies the best feasible outcomes were achieved. The number of FNs and FPs would be 0 in this case, and the total number of observed positive and negative classes would be the total number of TPs and TNs.

Tab. 3 shows the accuracy, sensitivity, specificity and error rate (%) of the VPS-PNN when applied to the 11 datasets, as well as the outcomes for the fundamental PNN and the findings given in the literature for the biogeography-based optimization (BBO) [51], the firefly algorithm (FA) [52], Enhanced-water cycle algorithm (E-WCA) [53] with PNN.

**Table 3:** Classification specificity, error rate, sensitivity, and accuracy for PNN, FA, BBO, E-WCA and VPS

| Dataset | Mechanism | TP | FP | TN | FN | Accuracy | Sensitivity | Specificity | Error rate |
|---|---|---|---|---|---|---|---|---|---|
| PID | PNN | 35 | 28 | 90 | 39 | 65.104 | 47.30 | 76.27 | 34.89 |
| | FA-PNN | 33 | 30 | 113 | 16 | 76.040 | *67.35* | 79.02 | 24.00 |
| | BBO-PNN | 38 | 25 | 99 | 30 | 71.350 | 55.88 | 79.84 | 28.56 |
| | E-WCA-PNN | 42 | 21 | 109 | 20 | 78.646 | 67.7 | 83.8 | 21.4 |
| | VPS-PNN | 39 | 24 | 122 | 7 | **84.884** | **85.71** | **84.48** | **15.12** |
| HSS | PNN | 44 | 12 | 6 | 15 | 64.93 | 74.58 | 33.33 | 35.07 |
| | FA-PNN | 54 | 2 | 10 | 11 | *83.12* | 83.08 | *83.33* | 16.88 |
| | BBO-PNN | 52 | 4 | 11 | 10 | 81.82 | *83.87* | 73.33 | 18.18 |
| | E-WCA-PNN | 53 | 3 | 11 | 10 | 83.12 | 84.10 | 78.60 | 16.90 |
| | VPS-PNN | 53 | 3 | 12 | 9 | **83.82** | **84.78** | **85.71** | **16.15** |
| AP | PNN | 23 | 1 | 1 | 2 | 88.88 | 92.00 | 50.00 | 11.12 |
| | FA-PNN | 24 | 0 | 1 | 2 | 92.59 | 92.31 | **100.00** | 7.41 |
| | BBO-PNN | 52 | 4 | 11 | 10 | 81.82 | *83.87* | 73.33 | 18.18 |
| | E-WCA-PNN | 24 | 0 | 2 | 1 | **96.29** | **96.00** | **100.00** | **3.70** |
| | VPS-PNN | 24 | 0 | 1 | 2 | 92.59 | 92.31 | **100.00** | 7.41 |

(Continued)

**Table 3 (continued)**

| Dataset | Mechanism | TP | FP | TN | FN | Accuracy | Sensitivity | Specificity | Error rate |
|---|---|---|---|---|---|---|---|---|---|
| BC | PNN | 14 | 9 | 36 | 13 | 69.44 | 51.9 | 80.00 | 30.60 |
| | FA-PNN | 31 | 1 | 24 | 12 | 80.88 | 72.09 | **96.00** | 19.12 |
| | BBO–PNN | 13 | 10 | 44 | 5 | 79.17 | 72.22 | 81.48 | 20.83 |
| | E-WCA-PNN | 14 | 9 | 47 | 2 | **84.72** | **87.50** | 83.90 | 15.30 |
| | VPS-PNN | 17 | 6 | 44 | 5 | **84.72** | 77.27 | 88.00 | **12.28** |
| LD | PNN | 18 | 15 | 34 | 19 | 60.46 | 48.60 | 69.40 | 39.50 |
| | FA-PNN | 31 | 1 | 24 | 12 | 79.07 | 72.09 | 96.00 | 20.93 |
| | BBO–PNN | 32 | 0 | 23 | 13 | 72.09 | 71.11 | **100.0** | 27.91 |
| | E-WCA-PNN | 23 | 10 | 45 | 8 | 79.06 | 74.20 | 81.80 | 20.90 |
| | VPS-PNN | 24 | 9 | 49 | 4 | **93.88** | **92.86** | **100.0** | **6.12** |
| Heart | PNN | 27 | 5 | 23 | 13 | 73.53 | 67.50 | 82.10 | 26.50 |
| | FA-PNN | 31 | 1 | 24 | 12 | 80.88 | 72.09 | 96.00 | 19.12 |
| | BBO–PNN | 32 | 0 | 23 | 13 | 80.90 | 71.11 | **100.0** | 19.10 |
| | E-WCA-PNN | 32 | 0 | 26 | 10 | **85.29** | 76.20 | **100.0** | **14.70** |
| | VPS-PNN | 164 | 15 | 43 | 28 | 82.80 | **85.05** | 75.00 | 17.20 |
| GCD | PNN | 133 | 46 | 39 | 32 | 68.80 | 80.60 | 45.90 | 31.20 |
| | FA-PNN | 166 | 13 | 30 | 41 | 78.40 | 80.19 | 69.77 | 21.60 |
| | BBO–PNN | 139 | 40 | 44 | 27 | 73.20 | 83.73 | 52.38 | 26.80 |
| | E-WCA-PNN | 178 | 3 | 16 | 55 | 76.8 | 76.20 | **84.20** | 23.2 |
| | VPS-PNN | 165 | 14 | 42 | 29 | **91.04** | **96.08** | 75.00 | **8.96** |
| Parkinson’s | PNN | 39 | 0 | 4 | 6 | 87.75 | 86.67 | 100.00 | 12.25 |
| | FA-PNN | 38 | 1 | 6 | 4 | 89.80 | 90.48 | 85.71 | 10.2 |
| | BBO–PNN | 39 | 0 | 7 | 3 | 93.88 | 92.86 | **100.00** | 6.12 |
| | E-WCA-PNN | 39 | 0 | 7 | 3 | 93.87 | 92.90 | **100.00** | 6.10 |
| | VPS-PNN | 71 | 3 | 94 | 5 | **95.38** | **93.42** | 96.91 | **4.62** |
| SPECTF | PNN | 49 | 4 | 5 | 9 | 80.59 | 84.48 | 55.56 | 19.40 |
| | FA-PNN | 52 | 1 | 10 | 4 | 92.54 | 92.86 | 90.91 | 7.46 |
| | BBO–PNN | 49 | 9 | 5 | 5 | 86.57 | 90.74 | 35.71 | 13.43 |
| | E-WCA-PNN | 52 | 1 | 13 | 1 | **97.01** | **98.10** | 92.9 | **3.0** |
| | VPS-PNN | 49 | 4 | 12 | 2 | 93.87 | 92.86 | **100.00** | 6.12 |
| ACA | PNN | 60 | 14 | 84 | 15 | 83.24 | 80.00 | 85.70 | 16.8 |
| | FA-PNN | 65 | 9 | 94 | 5 | 91.91 | 92.86 | 91.26 | 8.09 |
| | BBO-PNN | 65 | 9 | 88 | 11 | 88.53 | 85.53 | 90.72 | 11.47 |
| | E-WCA-PNN | 70 | 4 | 92 | 7 | 93.64 | 90.90 | 95.80 | 6.40 |
| | VPS-PNN | 71 | 3 | 94 | 5 | **95.38** | **93.42** | **96.91** | **4.62** |
| Fourclass | PNN | 59 | 19 | 127 | 11 | 86.11 | 84.29 | 86.99 | 13.89 |
| | FA-PNN | 78 | 0 | 138 | 0 | **100.00** | **100.00** | **100.00** | **0.00** |
| | BBO–PNN | 78 | 0 | 138 | 0 | **100.00** | **100.00** | **100.00** | **0.00** |
| | VPS-PNN | 78 | 0 | 138 | 0 | **100.00** | **100.00** | **100.00** | **0.00** |

For each dataset, the best accuracy outcome is highlighted in bold in Tab. 3. On all datasets, the VPS-PNN approach outperforms the PNN, FA, E-WCA and BBO methods based on the accuracy parameter.

## 9 Significance Testing

Based on the accuracy of classification, the T-test is utilized to evaluate the performance of the VPS method and FA. The T-test is utilized to see if the VPS algorithm outperforms the FA algorithm statistically with a significant interval of 95% ($\alpha = 0.05$) on classification accuracy. In Tab. 4, the mean, standard error mean and standard deviation, as well as the acquired p-values, are supplied as statistics relating to the correctness of the VPS algorithm and FA.

**Table 4:** T-test accuracy statistics and p-values for VPS with FA

| Dataset | | Mean | Sd. deviation | Sd. error mean | p-value |
|---|---|---|---|---|---|
| PID | VPS | 83.3333 | 2.23006 | 0.40720 | 0.00 |
| | FA | 73.4895 | 1.28560 | 0.23472 | |
| HSS | VPS | 83.1028 | 0.63377 | 0.11570 | 0.00 |
| | FA | 81.8179 | 1.02322 | 0.18681 | |
| AP | VPS | 92.5926 | 0.00000 | 0.00000 | 0.00 |
| | FA | 92.5926 | 0.00012 | 0.00002 | |
| BC | VPS | 82.9167 | 1.59613 | 0.29141 | 0.00 |
| | FA | 77.3935 | 1.74347 | 0.31831 | |
| LD | VPS | 92.5170 | 1.72282 | 0.31450 | 0.00 |
| | FA | 75.5810 | 1.49604 | 0.27310 | |
| Heart | VPS | 82.8000 | 0.00000 | 0.00000 | 0.00 |
| | FA | 78.6819 | 2.23781 | 0.40857 | |
| GCD | VPS | 88.2668 | 1.37125 | 0.25035 | 0.00 |
| | FA | 75.1600 | 1.58040 | 0.28854 | |
| Parkinson’s | VPS | 94.5087 | 0.75519 | 0.13788 | 0.00 |
| | FA | 89.7950 | 0.00000 | 0.00000 | |
| SPECTF | VPS | 92.5170 | 1.72282 | 0.31450 | 0.00 |
| | FA | 88.8057 | 1.82787 | 0.33372 | |
| ACA | VPS | 94.5087 | 0.75519 | 0.13788 | 0.00 |
| | FA | 89.8840 | 1.05983 | 0.19350 | |
| Fourclass | VPS | 100.000 | 0.00000 | 0.00000 | 0.00 |
| | FA | 100.000 | 0.00000 | 0.00000 | |

## 10 The Convergence Speed Results

When the VPS-PNN and FA-PNN the were used to the 11 datasets, Fig. 6 shows the simulation results of their convergence characteristics. Since the accuracy does not improve beyond the 180th iteration, each of the two models is subjected to a total of 200 iterations. For each of the 11 datasets, Fig. 6 depicts the convergence of the FA-PNN and VPS-PNN to the optimal classification accuracy. The y-axis is an accuracy and the x-axis is an iteration at Fig. 6. Fig. 7 presents box plots for the 11 datasets that show the VPS and FA's resolution quality dispersion.

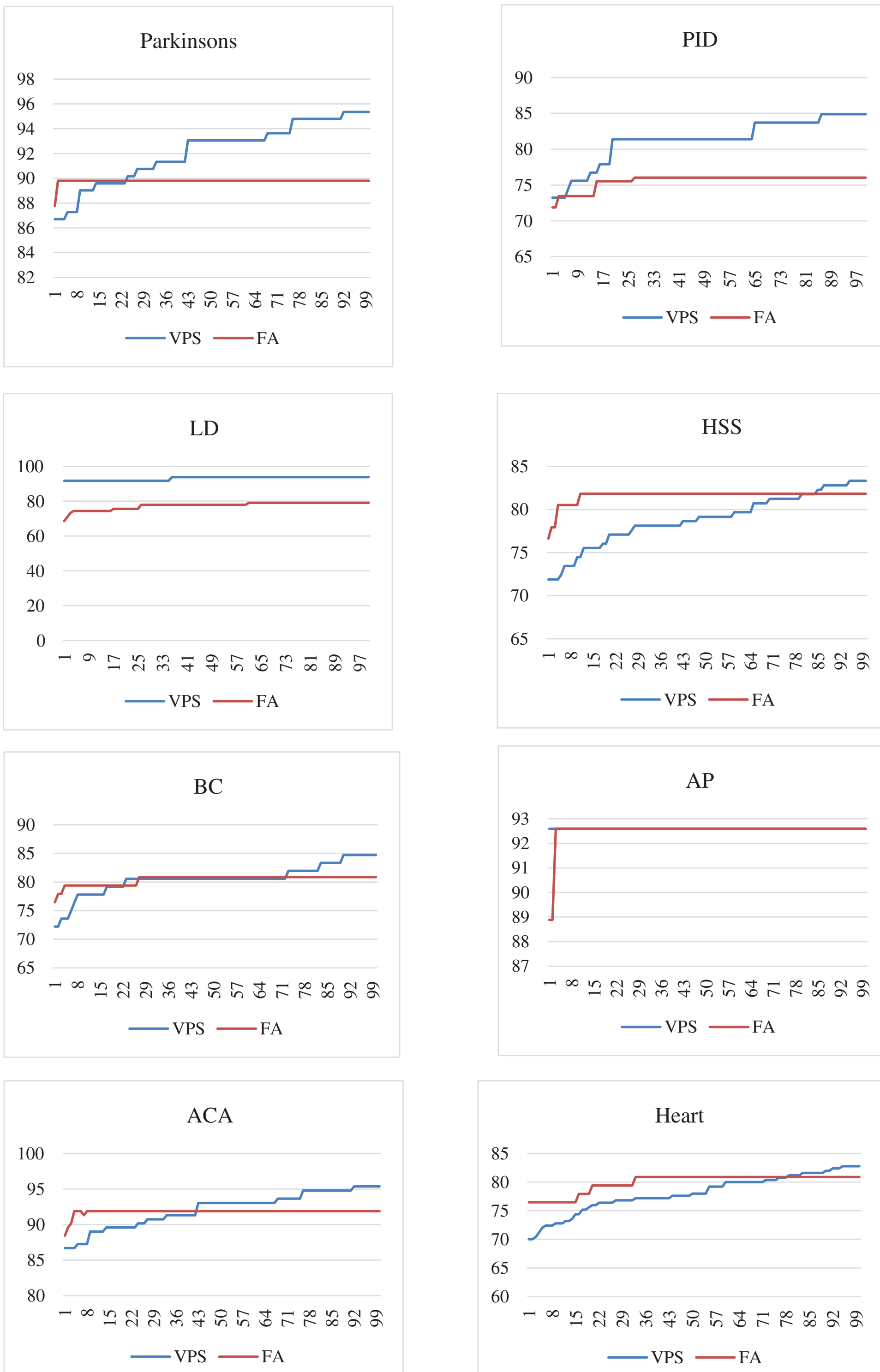


**Figure 6:** Continued

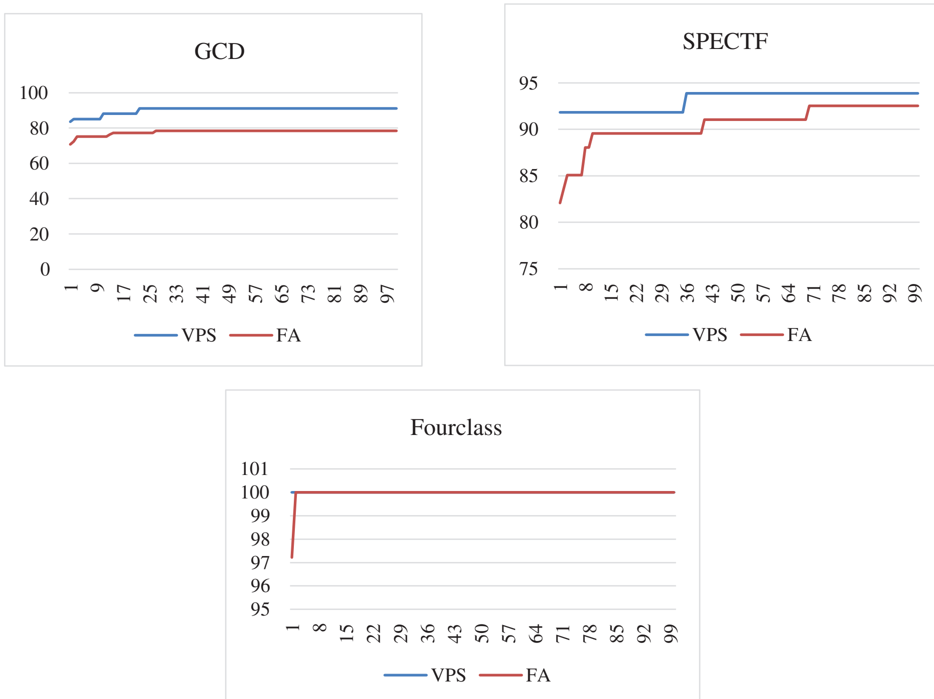


**Figure 6:** Convergence characteristics of VPS and FA

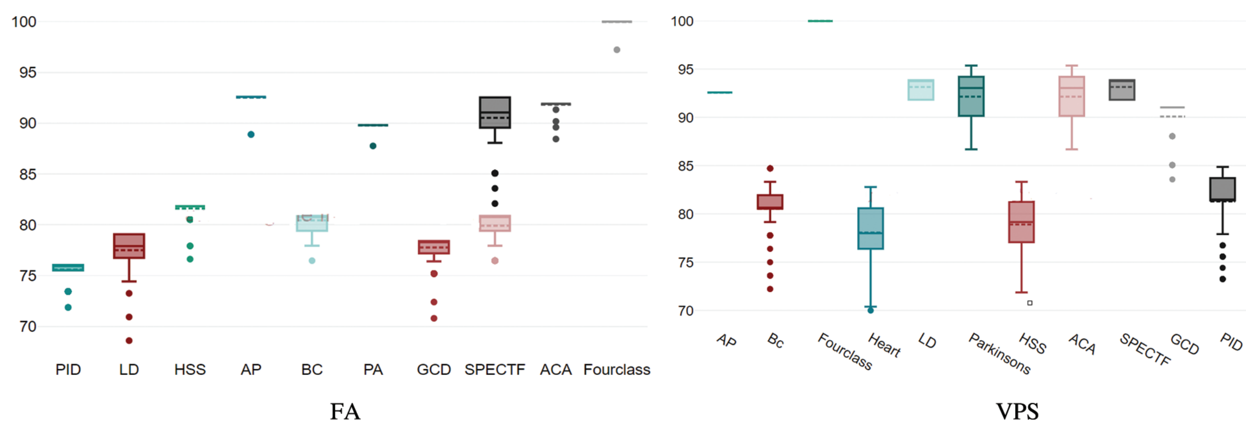


**Figure 7:** FA and VPS box plots

## 11 Conclusion

The authors of this study suggest the hybridization of VPS and PNN as a solution to classification difficulties. PNN was utilized to create the initial solutions randomly, which were subsequently optimized using the VPS. Experiments utilizing 11 benchmark datasets revealed that the suggested VPS-PNN outperformed the original PNN, FA, E-WCA and BBO on 8 out of 11 datasets. As a result, the VPS algorithm is an effective manner by efficiently exploring and exploiting the solution region in altering the weights of PNN for the purpose of assuring high accuracy in the mechanism of classification. Furthermore, we believe that more research is needed to determine how a good initial state can lead to higher classification accuracy for classification issues. In this respect, the enhanced vibrating particles system (EVPS) might be applied to other actual and high-dimensional datasets to investigate their behavior under various situations in terms of class and attribute numbers. As a result, this topic will be the focus of our future work.

**Funding Statement:** The authors received no specific funding for this study.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

[1] J. Han, M. Kamber and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed., no. 696, WS, Waltham, USA: Morgan Kaufmann Publishers, 2011.

[2] J. Manyika, M. Chui, B. Brown, J. Dobbs, C. Roxburgh *et al.,* "Big data: The next frontier for innovation, competition, and productivity," *McKinsey Global Institute*, vol. 5, no. 6, pp. 1–156. 2011.

[3] S. Agarwal, "Data mining: Data mining concepts and techniques," in *2013 Int. Conf. on Machine Intelligence and Research Advancement*, Katra India, IEEE, pp. 203–207, 2013.

[4] J. Sun, F. Khan, J. Li, M. D. Alshehri, R. Alturki *et al.,* "Mutual authentication scheme for the device-to-server communication in the internet of medical things," *IEEE Internet of Things Journal*, vol. 9, pp. 89344–89359, 2021.

[5] D. Li and Y. Du, *Artificial Intelligence with Uncertainty*, 2nd ed., no. 310, Boca Raton, Florida, USA, CRC Press, 2017.

[6] M. Alweshah, E. Ramadan, M. Ryalat, M. Almi'ani and A. Hammouri, "Water evaporation algorithm with probabilistic neural network for solving classification problems," *Jordanian Journal of Computers and Information Technology*, vol. 6, no. 1, pp. 1–9, 2021.

[7] R. L. Schalock, S. A. Borthwick, V. J. Bradley, W. H. Buntinx, D. L. Coulter *et al.,* "Intellectual disability: Definition, classification, and systems of supports," in *American Association on Intellectual and Developmental Disabilities*, 11th ed., vol. 26, no. 259, Washington, DC, AAIDD, 2010.

[8] 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," *Algorithms*, vol. 12, no. 8, pp. 160, 2019.

[9] M. Ibrahim, M. Wedyan, R. Alturki, M. A. Khan and A. Al-Jumaily, "Augmentation in healthcare: Augmented biosignal using deep learning and tensor representation," *Journal of Healthcare Engineering*, vol. 2021, pp. 1–9. 2021.

[10] A. A. Ibrahim and I. S. Yasseen, "Using neural networks to predict secondary structure for protein folding," *Journal of Computer and Communications*, vol. 5, no. 1, pp. 1–8, 2017.

[11] M. O. Wedyan, "Augmented reality and novel virtual sample generation algorithm based autism diagnosis system," Ph.D. dissertation, University of Technology Sydney, 2020.

[12] M. Alweshah, L. Rababa, M. H. Ryalat, A. Al Momani and M. F. Ababneh, "African buffalo algorithm: Training the probabilistic neural network to solve classification problems," *Journal of King Saud University-Computer and Information Sciences*, vol. 2020, pp. 1–11, 2020.

[13] M. Gendreau, M. Iori, G. Laporte and S. Martello, "A tabu search algorithm for a routing and container loading problem," *Transportation Science*, vol. 40, no. 3, pp. 342–350, 2006.

[14] K. A. Dowsland and J. J. Thompson, "Simulated annealing," in *Handbook of Natural Computing*, Berlin, Heidelberg, Springer-Verlag, pp. 1623–1655, 2012.

[15] E. H. L. Aarts and J. K. Lenstra, *Local Search in Combinatorial Optimization*, 2nd ed., no. 512, New Jersey, Princeton University Press, 2003.

[16] X. -S. Yang and X. He, "Firefly algorithm: Recent advances and applications," *International Journal of Swarm Intelligence*, vol. 1, no. 1, pp. 36–50, 2013.

[17] S. Mirjalili, "Genetic algorithm," in *Evolutionary Algorithms and Neural Networks*, vol. 780. Denmark, Cham, Springer, pp. 43–55, 2019.

[18] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. of ICNN'95-Int. Conf. on Neural Network*, Perth, WA, Australia, vol. 4, pp. 1942–1948, 1995.

[19] D. Karaboga, B. Gorkemli, C. Oztruk and N. Karaboga, "A comprehensive survey: Artificial bee colony (ABC) algorithm and applications," *Artificial Intelligence Review*, vol. 42, no. 1, pp. 21–57, 2014.

[20] S. Bhattacharya, P. K. R. Maddikunta, I. Meenakshisundaram, T. R. Gadekallu, S. Sharma *et al.,* "Deep neural networks based approach for battery life prediction," *Computers, Materials & Continua*, vol. 69, no. 2, pp. 2599–2615, 2021.

[21] S. Agrawal, S. Sarkar, G. Srivastava, P. K. R. Maddikunta and T. R. Gadekallu, "Genetically optimized prediction of remaining useful life," *Sustainable Computing: Informatics and Systems*, vol. 31, pp. 100565, 2021.

[22] T. R. Gadekallu and X. Z. Gao, "An efficient attribute reduction and fuzzy logic classifier for heart disease and diabetes prediction," *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science)*, vol. 14, no. 1, pp. 158–165, 2021.

[23] R. Kaluri, D. S. Rajput, Q. Xin, K. Lakshmanna, S. Bhattacharya *et al.,* "Roughsets-based approach for predicting battery life in IoT," arXiv preprint arXiv, 2021.

[24] X. Wu, X. Zhu, G. -Q. Wu and W. Ding, "Data mining with big data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 1, pp. 97–107, 2013.

[25] S. Madden, "From databases to big data," *IEEE Internet Computing*, vol. 16, no. 3, pp. 4–6, 2012.

[26] S. R. Pakize, "A comprehensive view of hadoop mapreduce scheduling algorithms," *International Journal of Computer Networks & Communications Security*, vol. 2, no. 9, pp. 308–317, 2014.

[27] S. Ghemawat, H. Gobioff and S. -T. Leung, "The google file system," in *Proc. of the Nineteenth ACM Symp. on Operating Systems Principles*, Bolton Landing NY USA, pp. 29–43, 2003.

[28] J. Dean and S. Ghemawat, "Mapreduce: A flexible data processing tool," *Communications of the ACM*, vol. 53, no. 1, pp. 72–77, 2010.

[29] C. Romero and S. Ventura, "Data mining in education," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 3.1, pp. 12–27, 2013.

[30] D. Che, M. Safran and Z. Peng, "From big data to big data mining: Challenges, issues, and opportunities," in *Int. Conf. on Database Systems for Advanced Applications*, Wuhan, China, Springer, pp. 1–15, 2013.

[31] H. A. Issad, R. Aoudjit and J. J. Rodrigues, "A comprehensive review of data mining techniques in smart agriculture," *Engineering in Agriculture, Environment and Food*, vol. 12, no. 4, pp. 511–525, 2019.

[32] W. Wange, J. Yang and R. Muntz, "STING: A statistical information grid approach to spatial data mining," *VLDB*, vol. 97, pp. 186–195, 1997.

[33] D. Pyle, *Data Preparation for Data Mining*, no. 560, San Francisco, CA, USA: Morgan Kaufmann, 1999.

[34] M. Herland, T. M. Khoshgoftaar and R. J. Wald, "A review of data mining using big data in health informatics," *Journal of Big Data*, vol. 1, no.1, pp. 1–35, 2014.

[35] Y. Zhang, S. -L. Guo, L. -N. Han and T. -L. Li, "Application and exploration of big data mining in clinical medicine," *Chinese Medical Journal*, vol. 129, no. 6, pp. 731, 2016.

[36] R. M. Cormack, "A review of classification," *Journal of the Royal Statistical Society: Series A (General)*, vol. 134, no. 3, pp. 321–353, 1971.

[37] S. Suthaharan, "Machine learning models and algorithms for big data classification," in *Integrated Series in Information Systems*, vol. 36, 1st ed., Boston, MA, USA: Springer, pp. 1–12, 2016.

[38] A. DeMaris, "A tutorial in logistic regression," *Journal of Marriage and the Family*, vol. 57, no. 4, pp. 956–968, 1995.

[39] M. J. L. Orr, *Introduction to Radial Basis Function Networks*, Scotland, UK: Center for Cognitive Science, Edinburgh University, 1996. [Online]. Available: http://anc.ed.ac.uk/rbf.

[40] E. Wiener, J. O. Pedersen and A. S. Weigend, "A neural network approach to topic spotting," in *Proc. of SDAIR-95, 4th Annual Symp. on Document Analysis and Information Retrieval*, Las Vegas, Nevada, Citeseer, vol. 317, pp. 332, 1995.

[41] L. Rokach and O. Maimon, "Clustering methods," in *Data Mining and Knowledge Discovery Handbook*, Springer: London, pp. 321–352, 2005.

[42] E. H. Ruspini, "A new approach to clustering," *Information and Control*, vol. 15, no. 1, pp. 22–32, 1969.

[43] R. Kaur, "A review of data mining techniques, applications and challenges," *International Journal of Multidisciplinary*, vol. 4, pp. 1–12, 2019.

[44] W. P. S. Jr, M. T. Musavi and J. N. Guidi, "Classification of chromosomes using a probabilistic neural network," *Cytometry: The Journal of the International Society for Analytical Cytology*, vol. 16, no. 1, pp. 17–24, 1994.

[45] D. F. Specht, "Probabilistic neural networks," *Neural Networks*, vol. 3, pp. 109–118, 1990.

[46] K. Z. Mao, K. -C. Tan and W. Ser, "Probabilistic neural-network structure determination for pattern classification," *IEEE Transactions on Neural Networks*, vol. 11, no. 4, pp. 1009–1016, 2000.

[47] P. D. Wasserman, *Advanced Methods in Neural Computing*, 605, Third Ave., New York, NY, United States: John Wiley & Sons, Inc., 1993.

[48] F. D. Specht, "Applications of probabilistic neural networks," *Applications of Artificial Neural Networks*, vol. 1294, pp. 344–353, 1990.

[49] A. Kaveh and M. I. Ghazaan, "A new meta-heuristic algorithm: Vibrating particles system," *Scientia Iranica. Transaction A, Civil Engineering*, vol. 24, no. 2, pp. 551–566, 2017.

[50] A. Kaveh and M. I. Ghazaan, "Matlab code for vibrating particles system algorithm," *Iran University of Science & Technology*, vol. 7, no. 3, pp. 355–366, 2017.

[51] M. Alweshah, A. I. Hammouri and S. Tedmori, "Biogeography-based optimisation for data classification problems," *International Journal of Data Mining, Modelling and Management*, vol. 9, no. 2, pp. 142–162, 2017.

[52] M. Alweshah and S. Abdullah, "Hybridizing firefly algorithms with a probabilistic neural network for solving classification problems," *Applied Soft Computing*, vol. 35, pp. 513–524, 2015.

[53] M. Alweshah, M. Al-Sendah, O. M. Dorgham, A. Al-Momani and S. Tedmori, "Improved water cycle algorithm with probabilistic neural network to solve classification problems," *Cluster Computing*, vol. 23, no. 4, pp. 2703–2718, 2020.