ORIGINAL ARTICLE

CrossMark

Construction biogeography-based optimization algorithm for solving classification problems

Mohammed Alweshah[1] (ID)

Abstract

Classification is a data mining task that assigns items in a collection to predefined categories or classes, also referred to as supervised learning. The goal of classification is to accurately predict the target class for each case in the data. A review of the literature shows that many algorithms, including statistical and machine learning algorithms, have been successfully used to handle classification problems in different areas, but their performance varies considerably. Even though the neural network is effective in addressing a wide range of problems, to date no specific neural network approach has been found that can ensure that the optimal solution is arrived at when solving classification problems. Some of the important challenges include finding the most appropriate weight parameter for the classifier through the implementation of population-based approaches; attaining a balance between the processes of exploration and exploitation by employing hybridization methods; and obtaining fast convergence by controlling random movement and by generating good initial solutions. This study investigates how can good initial populations drive higher convergence speed and better classification accuracy in solving classification problems. Local search (in this case, the simulated annealing algorithm) is used to produce an initial solution for the classification problem and then a heuristic initialization hybridized with biogeography-based optimization is applied. The proposed approaches are tested on 11 standard benchmark datasets. This is a new approach in the classification arena, and it represents an approach that outperforms the current state of the art on most of the tested benchmark datasets.

Keywords Data mining · Classification · Optimization · Metaheuristic · Biogeography-based algorithm

1 Introduction

Data mining refers to the procedure of uncovering unexpected patterns from huge volumes of data. Classification is a kind of data analysis that pulls out models that outline key data classes. These models, known as classifiers, project categorical (unordered, discrete) class labels. Classification is a key task in the data mining process. A training set is an input dataset within a classification problem, where the aim is to utilize the training dataset to construct a class label model. This model can then be employed to build up output when the intended output is not specified.

Several approaches, such as the neural network (NN) [1, 2], radial basis function (RBF) [3], Naive Bayes (NB) [4], A case-based reasoning [5], and support vector machine (SVM) [6, 7], have been effectively employed to solve classification problems [8]. Among these, the NN is a renowned and extensively deployed approach. Different NN models have been conceived, such as RBF, feed-forward, modular, multilayer perceptron (MLP), and probabilistic neural networks (PNNs).

Conversely, a metaheuristic methodology is an iterative generation procedure which steers a subordinate heuristic by blending intelligently dissimilar conceptions for delving into and utilizing the search space. Various methodologies have been deployed to categorize and define metaheuristic algorithms. However, essentially, both population and single based metaheuristics canister to train an NN. Techniques based on a single solution include the simulated annealing (SA) approach [9], iterated local search

✉ Mohammed Alweshah
weshah@bau.edu.jo

1 Department of Computer Science, Prince Abdullah Bin Ghazi Faculty of Information Technology, Al-Balqa Applied University, Salt 19117, Jordan

approach [10], Tabu search [11], guided local search [12], and variable neighborhood search [13].

The utilization of a population-based technique along with an NN has drawn lot of attention. The reason behind this is that when an NN is blended with an evolutionary algorithm (EA), a system with greater intelligence can be developed than when only an EA or an NN is utilized [14]. For example, the particle swarm optimization (PSO) algorithm has been utilized for training an NN [15]. Other swarm intelligence approaches include ant colony optimization (ACO) [16], the genetic algorithm (GA) [17], artificial bee colony (ABC) algorithm [18], harmony search algorithm (HSA) [19], cuckoo search (CS) algorithm [20], artificial fish swarm (AFS) algorithm [21], firefly algorithm (FA) with PNN [22, 23], and biogeography-based optimization (BBO) [24].

BBO is a newly created population-based stochastic metaheuristic optimization algorithm that was first introduced in 2008 by Dan Simon [25]. The performance of BBO has been assessed by comparing it with seven other biology-based algorithms. The results obtained are strongly in favor of the BBO algorithm. It has been applied in several domains and fields, such as image processing, scheduling, image classification, and feature extraction. Moreover, the efficiency of the BBO algorithm has been demonstrated when applied to a real-world sensor selection problem for aircraft engine health estimation [26]. The performance of BBO has been compared with many algorithms in the literature. It has performed well, which has encouraged other researchers to improve the BBO algorithm by using it in conjunction with different methods in order to solve different optimization problems.

Nevertheless, despite recent advances there is still a need for more research to determine how a sound initial state can steer an algorithm toward improved classification precision and greater speed of convergence when solving classification problems. Therefore, this study deploys local search (i.e., the SA algorithm) to formulate an initial solution for classification problems. It then employs a heuristic initialization hybridized with BBO to enhance the quality of the initial solution by painstakingly probing around it, thereby enhancing the solutions' performance in determining the best weight for the PNN classification approach. The research presented investigates how the initial population of BBO can be generated in order to hasten convergence and the ultimate objective values.

The remainder of this research paper organized as follows: In Sect. 2 explains the initialization of the population. Section 3 explains the hybridized heuristic initialization utilizing the proposed initialization of biogeography-based optimization (IBBO) algorithm. Section 4 describes and analyzes the results of an experiment that tests the IBBO algorithm on 11 benchmark datasets. Finally, Sect. 5 concludes the paper.

2 Initialization of the population

The techniques deployed for population initialization can be categorized into sequential diversification, random generation, heuristic initialization, and parallel diversification [25]. The initial population of a metaheuristic algorithm is produced randomly. Any heuristic (e.g., local search) can be utilized to initialize the population [25]. Heuristic initialization begins from scratch (i.e., from a null/empty solution) and builds a solution by handing over values to one decision variable at a time until an entire solution is produced. At every step, a local search heuristic is employed to choose a new solution, and this approach seems to be more effective compared to random initialization [25].

Local search or any other metaheuristic algorithm can be used to generate the initial population of an algorithm. In this research, the SA algorithm [26] is used to generate the initial population by generating a random solution (as mentioned in the random construction method) and then applying the SA algorithm for a certain number of iterations to improve the quality of the solution. The improved solution is then added to the algorithm's population. These steps are repeated until the population is filled with the required number of solutions. However, it should be noted that although a local search method can generate an initial population that contains a good-quality solution, the population may lose its diversity which may lead to premature convergence. A general pseudocode of the local search constructive heuristic is shown in Fig. 1.

3 Proposed method: initial population of BBO algorithm

Biogeography-based optimization has been applied in several domains and fields, such as image processing, scheduling, image classification, and feature extraction [27]. The results obtained by the BBO algorithm compare favorably with some other classical metaheuristics such as the GA, PSO, and ACO when applied to a similar problem,

```
Population  ←  Create empty population.
For i = 1 to Ps          // Ps is the desired population size
        S ← new random solution
        Repeat
            R ← Copy(S)
            Perturb(R)        // Randomly modify solution R
            If Quality(R) is better than Quality(S) then
                S ← R
            End if
        Until stopping criterion is satisfied
        Population ← Population ∪ S
End For
```

Fig. 1 Pseudocode for local search constructive heuristic

and the BBO algorithm also consistently outperforms these algorithms in finding an optimal solution for some other problems [27]. Moreover, experimental results have revealed that the BBO algorithm can be considered a promising alternative approach for solving optimal power flow problems because it can converge to a better-quality solution [24]. It also has good convergence characteristics and robustness compared with PSO, differential evolution (DE), and some other approaches [27].

In recent years, BBO has received greater attention due to its uniqueness and amazing performance. Also, it performs well when compared with some other metaheuristics such as the GA, PSO, and ACO [27]. Also, they have confirmed their influential skills in simulations and experimentations of computation and optimization. Biogeography-based optimization can show significant characters in numerous areas, including data mining, bioinformatics, and information security, and its current applications indicate that BBO has the capacity for real-world deployment [27].

BBO has been proved to be very efficient in solving many NP hard problems, that is the problems for which event the best-known algorithms have exponential time complexity [28]. Typically, the time complexity of which is O(n). It is not difficult to see that the complexity of each iteration of the algorithm is $O(n^2 \ D + nO(f))$, where $O(f)$ is the time complexity for computing the fitness function f [29].

Therefore, the aim of this study is to investigate the impact of a good initial population on the convergence speed and the final objective values by using heuristic initialization (i.e., local search) to help the BBO algorithm to create the initial population. Local search (i.e., the SA algorithm) is used to create a good initial population in order to reduce the number of iterations (i.e., to achieve a better convergence speed) and to enhance the quality of the initial population so that high-quality solutions can be found.

Specifically, this study proposes a sequential hybridization of SA and BBO rather than the version of BBO applied in Mirjalili [30] and others. Basically, in this study, it is hypothesized that having a good initial point for BBO will ameliorate the finding of the optimal values for NNs. Thus, the aim of this study is to use a heuristic initialization (i.e., local search) to help the BBO in determining the initial population to see the good solution from anywhere in the problem space and prevent corresponds as a random search. The local search (i.e., the SA algorithm) is used to create a good initial population to reduce the number of iterations in BBO (i.e., to achieve a better convergence speed) and to enhance the quality of the initial population so that high-quality solutions can be found.

Figures 2 and 3 show the proposed structure of the IBBO algorithm. Figure 2 shows the first part of the proposed algorithm which involves using the SA algorithm to produce different initial solutions (hybrid initialization of the population) to find a good initial population state that could serve to maximize classification accuracy also and thus overcome the low convergence issue when solving classification problems. The SA algorithm is based on the ideas of statistical mechanics whereby the annealing process is used to heat a substance which is then slowly cooled to create a powerful crystalline structure. The strength of the construction is dependent upon the speed of the cooling of the metals. The SA algorithm was the first algorithm accepting non-improving neighbors to be selected. SA escaping from local optima and the better transfer are all the time accepted [31].

The second part of the proposed approach, which is shown in Fig. 3, involves the BBO, which is used as an improvement algorithm that aims to try to control random steps in the BBO mechanism, after evaluating the population in each iteration. The BBO approach is one of the efficient methods for solving complex problems, so BBO was employed to have optimal values for the optimize weight of the PNN [24].

Thus, the IBBO algorithm is composed of two main stages, namely initialization (SA) and BBO. A set of steps is performed in each of these stages. In the SA stage, the SA algorithm is used to make the initial (solution) population of candidate solutions for the weights of standard PNN (in this case). In the second stage, the fitness value of all the candidates is calculated according to the BBO algorithm and searches are conducted to find better solutions to ensure efficient convergence, quality solutions, and ultimately obtain the optimal parameter settings for training the PNN, and thereby achieve better classification accuracy.

4 Experimental results

An experiment was conducted to assess and compare the performance of the proposed IBBO approach and that of some other hybrid approaches in terms of classification quality and convergence speed when applied to a number of datasets. The results for classification quality are presented in tabular form in terms of accuracy, specificity, sensitivity, and error rate. The results for convergence speed are presented as graphical depictions of the algorithms' convergence behavior. The results thus obtained are corroborated by employing a statistical significance test (i.e., the Wilcoxon test), and a brief commentary is provided.

The experiment was performed on 11 datasets that were used in another study [22]. The size of the training and testing sets was also adopted from that study. Table 1
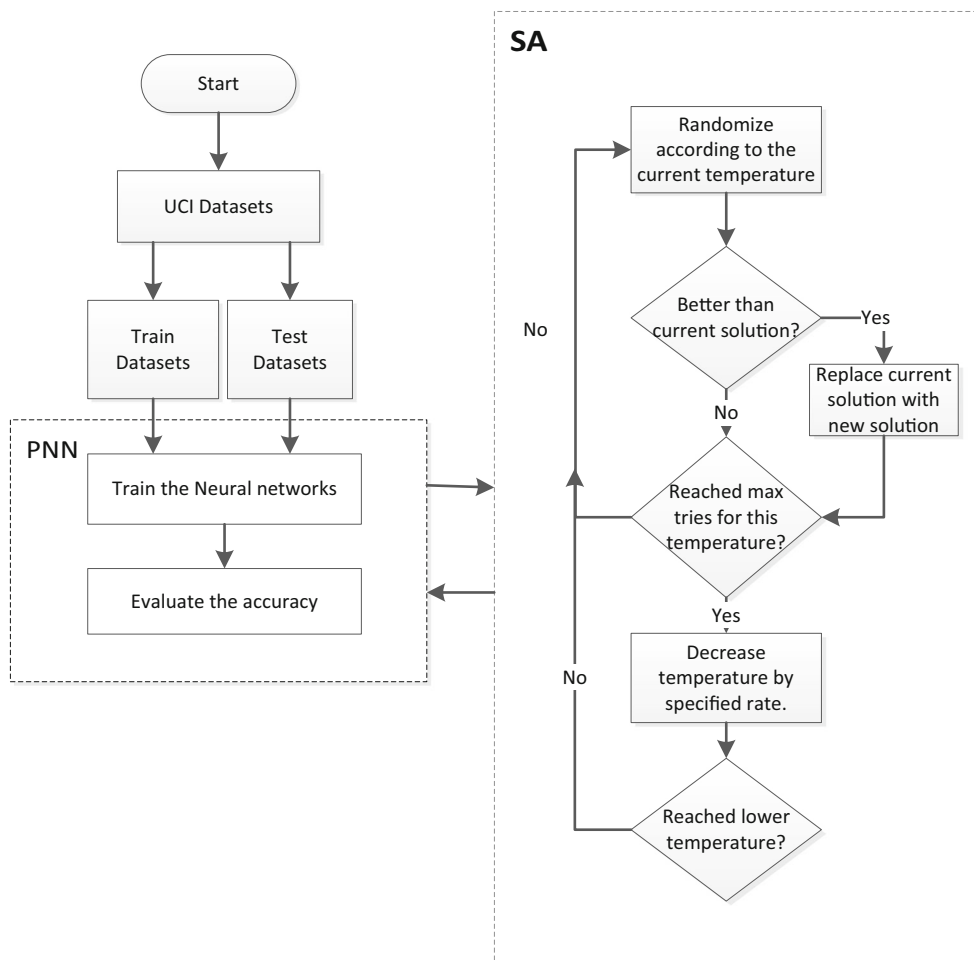
Fig. 2 Flowchart of the first stage of the proposed IBBO approach: heuristic initialization by SA

summarizes the characteristics of the datasets, and Table 2 summarizes the parameters for the proposed algorithm that were determined after number of preliminary experiments.

4.1 Results for classification quality

The classification quality achieved by the tested algorithms depended on their generation of a sound initial population state for facilitating faster convergence and enhancing solution quality. Table 3 provides a comparison of the performance of the proposed IBBO and that of BBO when a preliminary solution is randomly generated. The optimal results are indicated in bold.

The first contribution of this study is the proposition of a new hybrid method that combines the PNN and BBO to solve classification problems using random initialization. The second contribution is the use of a heuristic initialization (i.e., local search) in order to help the BBO to start from a good solution. Local search (i.e., the SA algorithm) is used to create a good initial population in order to reduce the number of iterations of BBO and thus enhance the

quality of the solutions. Table 3 shows a comparison of the results obtained by IBBO with PNN and those achieved by BBO-PNN with a randomly generated preliminary solution.

As can be seen from Table 3, the hybrid IBBO methodology shows improved performance with regard to classification precision compared to the BBO. The IBBO clearly outshines the BBO technique and can attain greatest classification precision for the tested datasets. Table 3 also indicates that the IBBO shows better performance with depend on sensitivity than the BBO on the all of the datasets. For instance, in the case of the BC dataset, IBBO attained a sensitivity of 83%. In other words, when a diagnostic test is conducted to find a patient who has a particular disease, there is an 83% prospect of positively detecting the patient. As for specificity, in the case of the ACA dataset, IBBO attained 98%. This means that if a diagnostic test is conducted on a patient who does not have a specific disease, there is a 98% chance of detecting a patient who is a negative case. In light of the above-reported results, the IBBO is superior to the BBO with regard
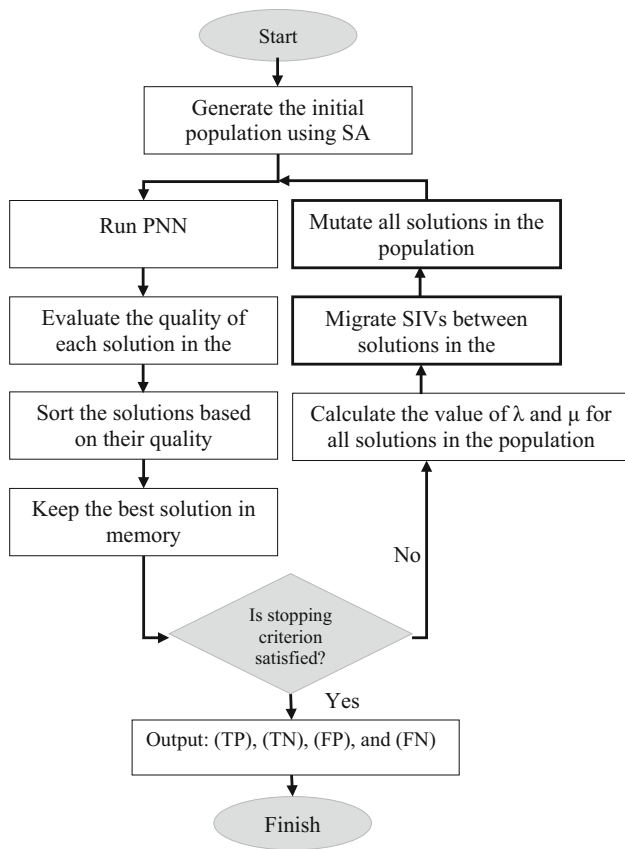
Fig. 3 Flowchart of the second stage of the proposed IBBO approach

to classification precision, specificity, sensitivity, and error rate. This is because when the algorithm begins from a good initial state, it can improve the ultimate objective values.

4.2 Results for convergence speed

The convergence behavior of the proposed IBBO algorithm was examined with regard to the generation of various

Table 2 Parameter settings

Parameter	Value
Population size	50
Number of generations/iterations	100
Mutation Rate (μ)	0.005
Initial temperature T_0	100
Final temperature T_f	0.5

preliminary solutions for determining a good initial population state so as to warrant effectual convergence, centered around the BBO on the test datasets.

Figure 4 depicts the simulation outcomes of the convergence attributes of the IBBO algorithm and that of another BBO technique, the (LSFA) [22]. The figure shows that IBBO has a faster convergence than LSFA and that the convergence trend is the same for all the datasets tested.

As the simulation results indicate, the search scheme of the IBBO algorithm accelerates convergence by blending the heuristic initialization approach with the LSFA. Prior research works have observed that heuristic initialization can be employed to accelerate the search process and can also be utilized to attain an almost-optimal solution [32]. Here, the heuristic initialization algorithm is able to come up with viable initial solutions, thereby driving quicker convergence. According to the simulation outcomes, the IBBO exhibits quicker convergence in comparison with other models primarily because this algorithm begins after a good preliminary population state.

In sum, IBBO is an appropriate approach that can be deployed to solve classification problems, considering that it exhibited decent performance with regard to both classification quality and speed of convergence.

Table 1 Datasets characteristics

Dataset	Number. of attributes	Training data	Testing data	No. of classes
PIMA Indian diabetes dataset (PID)	8	518	192	2
Dataset of Haberman surgery survival (HSS)	3	206	77	2
The appendicitis dataset (AP)	7	71	27	2
Breast cancer datasets (BC)	10	193	72	2
Liver disorders (LD)	6	233	86	2
Heart dataset (Heart)	13	182	68	2
German credit data set (GCD)	20	675	250	2
Parkinson's	23	131	49	2
SPECTF	45	180	67	2
Australian credit approval (ACA)	14	465	173	2
Fourclass	2	581	216	2

Table 3 Classification accuracy, error rate, sensitivity, and specificity (%) for the IBBO algorithm and BBO-PNN

	Approach	TP.	FP.	TN.	FN.	Accuracy	Sensitivity	Specificity	Error rate
PID	BBO	50	13	123	6	90.10	0.79	**0.95**	0.10
	IBBO	54	9	121	8	**91.15**	**0.86**	0.94	**0.09**
HSS	BBO	53	3	13	8	85.71	0.95	0.62	0.14
	IBBO	56	0	14	7	**90.9**	**1.00**	**0.67**	**0.09**
AP	BBO	24	0	1	2	92.59	1.00	0.33	0.07
	IBBO	24	0	2	1	**96.3**	**1.00**	**0.67**	**0.04**
BC	BBO	16	7	45	4	84.72	0.70	**0.92**	0.15
	IBBO	19	4	44	5	**87.5**	**0.83**	0.90	**0.13**
LD	BBO	27	6	48	0	87.21	0.82	**1.00**	0.07
	IBBO	31	2	50	3	**94.19**	**0.94**	0.94	**0.06**
Heart	BBO	32	0	24	12	82.35	1.00	0.67	0.18
	IBBO	32	0	27	9	**86.76**	**1.00**	**0.75**	**0.13**
GCD	BBO	169	10	53	18	88.80	0.94	0.75	0.11
	IBBO	169	10	55	16	**89.60**	**0.94**	**0.77**	**0.10**
Parkinson's	BBO	39	0	7	3	93.88	1.00	0.70	0.06
	IBBO	39	0	8	2	**95.92**	**1.00**	**0.80**	**0.04**
SPECTF	BBO	49	4	14	0	94.03	0.92	**1.00**	0.06
	IBBO	53	0	11	3	**95.52**	**1.00**	0.79	**0.04**
ACA	BBO	73	1	94	5	96.53	**0.99**	0.95	0.03
	IBBO	72	2	97	2	**97.69**	0.97	**0.98**	**0.02**
Fourclass	BBO	78	0	138	0	100.00	1.00	1.00	0.00
	IBBO	78	0	138	0	**100.00**	**1.00**	**1.00**	**0.00**

The optimal results are indicated in bold

4.3 Comparison with the literature

The proposed IBBO approach is also compared with some renowned approaches in the literature which have attained the best-known outcomes. The approaches used for this comparison are neural fuzzy inference system [24], fuzzy NNs [25], back-propagation (BP) [32], artificial neural networks (ANNs) [33] (CBA) [34], and hybrid simplex-GA [35], as well as the LSFA, SFA, LFA, and FA [22]. The acronyms for these state-of-the-art methods are given in Table 4.

The outcomes of the comparison are based on 20 independent runs on 11 benchmark datasets. The optimal outcomes are depicted in bold in Table 5. It should be noted that none of the methodologies in Table 4 tested on wholly datasets utilized in this study. Therefore, the comparison of the performance of the proposed approach differs by dataset in terms of the methodologies compared and depends on various parameter settings which could lead to different classification precision results (Table 3).

As can be seen from Tables 5 and 6, IBBO performs best overall, followed by LSFA in terms of classification quality. Furthermore, IBBO is able to categorize the Fourclass dataset, sans any erroneous classification (precision of 100%). The outcomes of the experiment, as depicted in Tables 5 and 6, suggest that the proposed

hybrid method (IBBO) outclasses the other methodologies in the literature on seven out of the 11 datasets (HSS, AP, Heart, BC, SPECTF, ACA and Fourclass). Furthermore, IBBO exhibits superior performance with regards to classification precision compared to BBO. The IBBO undoubtedly outclassed the other hybrid approaches and can attain the highest classification accuracy on all the datasets tested.

4.4 Validation and analysis of results

A statistical analysis of the performance of IBBO was also undertaken to ascertain whether it was statistically dissimilar from the LSFA. For this purpose, a Wilcoxon test was conducted with a significantly of 95% ($\alpha = 0.05$) for classification precision, specificity, and sensitivity. Table 7 depicts the p values obtained.

From Table 6, it can be seen that, with regard to the classification precision, the outcomes attained by the IBBO are statistically different with p value less than 0.05) compared to those of the LSFA algorithm (with the exception of the ACA, GCD and Fourclass datasets). Notably, there was no significant difference when using LSFA and IBBO algorithms in six datasets that exhibited a p value more than 0.05 (indicated in bold). With regard to
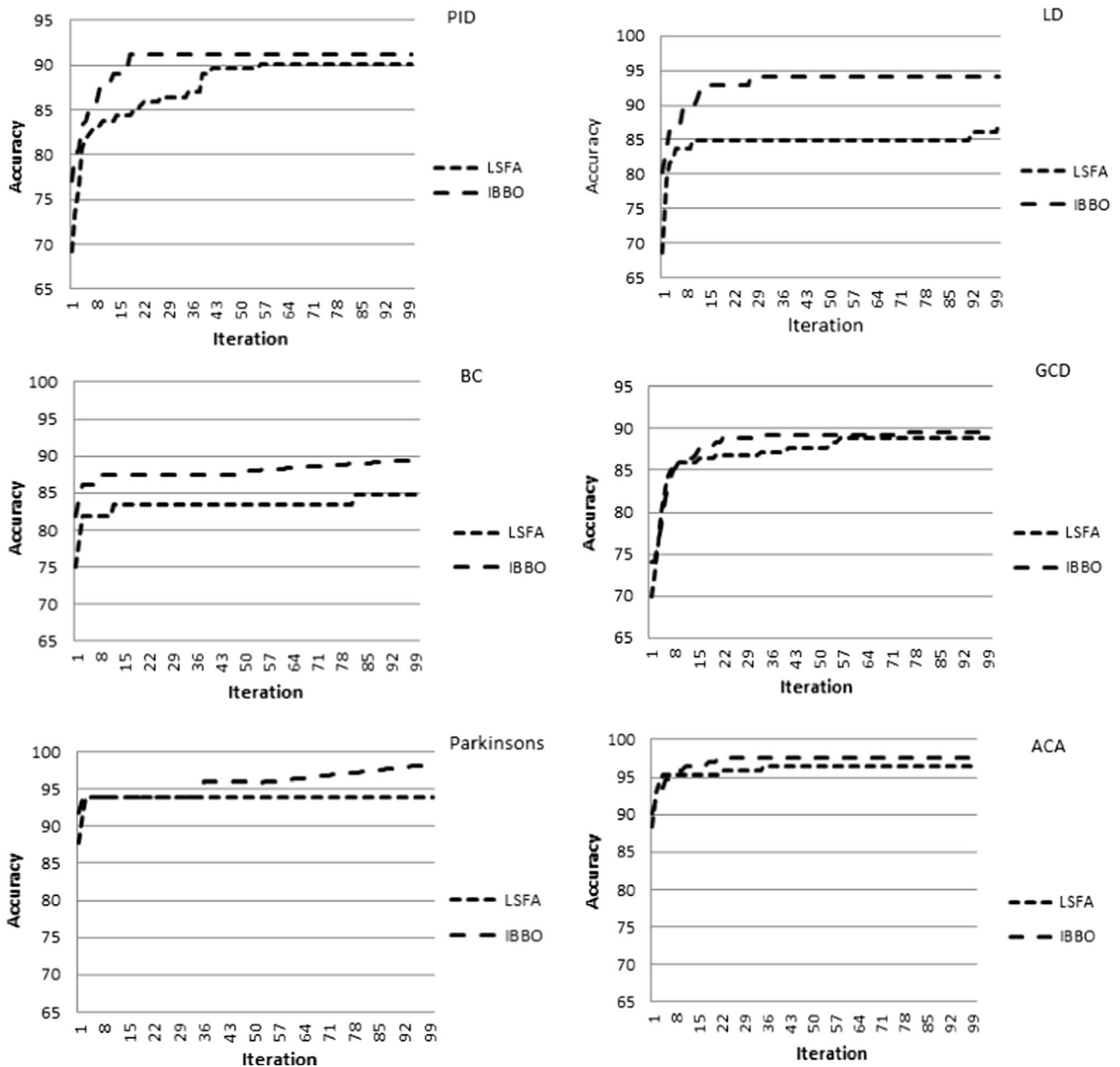
Fig. 4 Convergence characteristics of LSFA and IBBO algorithms

Table 4 Acronyms of the compared methods

#	Acronym	Name of approach	References
1	M_1	Neural fuzzy inference system	[33]
2	M_2	Fuzzy NNs	[36]
3	M_3	BP	[37]
4	M_4	ANNs	[38]
5	M_5	CBA	[39]
6	M_6	LSFA, SFA, LFA, FA	[22]

specificity, the IBBO and the LSFA are statistically different (except for the ACA, LD and Fourclass datasets).

One of the main shortcomings of the BBO can be found on the exploitation side, and this shortcoming led researchers to propose modified versions of BBO [27]. One of the main reasons for proposing the hybrid IBBO approach in this study was to increase the exploitation capability by using local search (i.e., the SA algorithm) to allow BBO to be initiated with a strong population.

Table 5 Comparison between the results using *IBBO* and state of the art

Dataset	Algorithms	Classification accuracy (%)
PID	IBBO	**91.15** (First rank)
	LSFA	90.10
	SFA	88.54
	LFA	77.08
	FA	76.04
	M1	78.6
	M2	81.8
	M5	76.7
HSS	IBBO	**90.9** (First rank)
	LSFA	85.71
	SFA	84.41
	LFA	83.12
	FA	83.11
	M5	72.7
AP	IBBO	**96.3** (First rank)
	LSFA	92.59
	SFA	92.59
	LFA	92.59
	FA	92.59
	M5	91.4
BC	IBBO	**87.5** (First rank)
	LSFA	84.72
	SFA	84.00
	LFA	81.94
	FA	80.88
	M8	73.5
	M16	79.687
LD	IBBO	**94.19** (First rank)
	LSFA	87.21
	SFA	86.05
	LFA	82.56
	FA	79.07
	M5	77.5
Heart	IBBO	**86.76** (First rank)
	LSFA	82.35
	SFA	82.35
	LFA	80.88
	FA	80.88
	M5	77.9
GCD	IBBO	**88.80** (First rank)
	LSFA	**88.80** (First rank)
	SFA	87.60
	LFA	78.40
	FA	78.40
	M5	86.4

The optimal results are indicated in bold

Table 6 Comparison of results using the second part of the datasets between *IBBO* and state of the art

Parkinson's	IBBO	**95.92** (First rank)
	LSFA	93.88
	SFA	91.84
	LFA	89.80
	FA	89.80
	M4	81.3
	M5	85.7
SPECTF	IBBO	**95.52** (First rank)
	LSFA	94.03
	SFA	94.03
	LFA	92.54
	FA	92.54
	M5	83.6
ACA	IBBO	**97.69** (First rank)
	LSFA	96.53
	SFA	95.37
	LFA	91.91
	FA	91.91
	M5	69.9
Fourclass	IBBO	**100** (First rank)
	LSFA	**100** (First rank)
	SFA	**100** (First rank)
	LFA	**100** (First rank)
	FA	**100** (First rank)
	M5	94.6

The optimal results are indicated in bold

Table 7 *p* values using Wilcoxon test for the algorithms of IBBO and LSFA for classification accuracy, sensitivity and specificity

Dataset	IBBO versus LSFA		
	Accuracy	Sensitivity	Specificity
PID	0.022	0.005	0.005
HSS	0.005	0.010	0.013
AP	0.003	**1.000**	0.005
BC	0.005	**0.504**	0.007
LD	0.005	0.008	**0.121**
Heart	0.004	**1.000**	0.004
GCD	**0.774**	0.044	0.032
Parkinson's	0.046	**1.000**	0.046
SPECTF	0.058	0.005	0.012
ACA	**0.550**	**0.062**	**0.416**
Fourclass	**1.000**	**1.000**	**1.000**

No significant difference is indicated in bold

5 Conclusion

The ultimate aim of this study was to enhance the solutions generated for classification problems by addressing two challenges: (1) making adjustments to the internal weights of the network to enable high-quality classification precision and (2) attaining a faster convergence speed. This study examined the feasibility of using BBO along with a PNN algorithm. Here, the BBO was utilized to make adjustments to the PNN's weights. Then, additional research was carried out to ascertain how an initial state could determine improved classification precision with faster convergence speed when resolving classification problems. This study employed the SA algorithm to produce a preliminary result for the improve classification problem. Then BBO was applied to enhance the quality of the preliminary solution by painstakingly searching around it to determine the best weight for enhanced PNN classification methodology. The proposed hybridized IBBO methodology exhibited superior overall the performance of 11 benchmark datasets used. Going forward, the proposed approach could be further developed and applied to real-value datasets in various domains and also to major real-world problems including text and web mining. Recently, other advanced versions of BBO have been proposed for highly demand none convex optimization problems. These versions could be tailored for classification problems and then compared with the IBBO proposed in this study.

Compliance with ethical standards

Conflict of interest The author certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

References

1. Specht DF (1991) A general regression neural network. IEEE Trans Neural Netw 2:568–576
2. Alweshah M, Omar A, Alzubi J, Alaqeel S (2016) Solving attribute reduction problem using wrapper genetic programming. Int J Comput Sci Netw Secur 16:77–84
3. Warwick K, Craddock R (1996) An introduction to radial basis functions for system identification. A comparison with other neural network methods. In: Proceedings of the 35th IEEE conference on decision and control, vol 1, pp 464–469
4. Friedman N, Geiger D, Goldszmidt M (1997) Bayesian network classifiers. Mach Learn 29:131–163
5. Alshareef AM, Bakar AA, Hamdan AR, Abdullah SMS, Alweshah M (2015) A case-based reasoning approach for pattern detection in Malaysia rainfall data. Int J Big Data Intell 2:285–302
6. Han J, Kamber M, Pei J (2011) Data mining: concepts and techniques. Elsevier, Amsterdam
7. Alweshah M, Hammouri AI, Rashaideh H, Ababneh M, Tayyeb H (2017) Solving time series classification problems using combined of support vector machine and neural network. In: International journal of data analysis techniques and strategies (**in press**)
8. Alshareef A, Ahmida S, Bakar AA, Hamdan AR, Alweshah M (2015) Mining survey data on university students to determine trends in the selection of majors. Sci Inf Conf (SAI) 2015:586–590
9. Ren K, Qu J (2014) Identification of shaft centerline orbit for wind power units based on Hopfield neural network improved by simulated annealing. In: Mathematical problems in engineering
10. El-Bouri A (2012) An investigation of initial solutions on the performance of an iterated local search algorithm for the permutation flowshop. In: 2012 IEEE congress on evolutionary computation (CEC), pp 1–5
11. Syed Mustafa A, Kumara Swamy YS (2015) Web service classification using multi-layer perceptron optimized with Tabu search. In: 2015 IEEE international on advance computing conference (IACC), pp 290–294
12. Tairan N, Qingfu Z (2010) Population-based guided local search: some preliminary experimental results. In: 2010 IEEE congress on evolutionary computation (CEC), pp 1–5
13. Xie X, Huibo Z, Yanping L, Yongyue Z (2012) Variable neighborhood search based multi-objective dynamic crane scheduling. In: 2012 international conference on measurement, information and control (MIC), pp 457–460
14. Malinak P, Jaksa R (2007) Simultaneous gradient and evolutionary neural network weights adaptation methods. In: IEEE congress on evolutionary computation, CEC, pp 2665–2671
15. Ekkachai K, Nilkhamhang I (2016) Swing phase control of semi-active prosthetic knee using neural network predictive control with particle swarm optimization. IEEE Trans Neural Syst Rehabil Eng 24:1169
16. Fei H, Dan Y, Qing-Hua L, De-Shuang H (2015) A novel diversity-guided ensemble of neural network based on attractive and repulsive particle swarm optimization. In: 2015 international joint conference neural networks (IJCNN), pp 1–7
17. Hu L, Qin L, Mao K, Chen W, Fu X (2016) Optimization of neural network by genetic algorithm for flowrate determination in multipath ultrasonic gas flowmeter. Sens J EEE 16:1158–1167
18. Dandan L, Wanxin X, Yilei P (2015) A high-precision prediction model using Ant Colony Algorithm and neural network, In: 2015 international conference on logistics, informatics and service sciences (LISS), pp 1–6
19. Ilkucar M, Isik AH, Cifci A (2014) Classification of breast cancer data with harmony search and back propagation based artificial neural network. In: 2014 22nd signal processing and communications applications conference (SIU), pp 762–765
20. Xiaojin X, Yun P, Ruijuan J, Yilan L (2015) Optimizing neural network classification by using the Cuckoo algorithm. In: 2015 11th international conference on natural computation (ICNC), pp 24–30
21. Mengmeng L, Li T, Li Z, Dabo Z (2014) A fault diagnosis algorithm using probabilistic neural network with particle fish swarm algorithm. In: 2014 11th World Congress on, intelligent control and automation (WCICA), pp 3713–3717
22. Alweshah M, Abdullah S (2015) Hybridizing firefly algorithms with a probabilistic neural network for solving classification problems. Appl Soft Comput 35:513–524

23. Alweshah M (2014) Firefly algorithm with artificial neural network for time series problems. Res J Appl Sci Eng Technol 7:3978–3982

24. Alweshah M, Hammouri AI, Tedmori S (2017) Biogeography-based optimisation for data classification problems. Int J Data Min Modelling Manag 9:142–162

25. Talbi E-G (2009) Metaheuristics: from design to implementation. Wiley, Hoboken

26. S. Sakamoto, E. Kulla, T. Oda, M. Ikeda, L. Barolli, and F. Xhafa, "A comparison study of Hill Climbing, Simulated Annealing and Genetic Algorithm for node placement problem in WMNs," *Journal of High Speed Networks,* vol. 20, pp. 55-66, 01/01/2014

27. Guo W, Chen M, Wang L, Mao Y, Wu Q (2017) A survey of biogeography-based optimization. Neural Comput Appl 28:1909–1926

28. Ammu P, Sivakumar K, Rejimoan R (2013) Biogeography-based optimization-a survey. Int J Electron Comput Sci Eng 2:154–160

29. Wang Z-C, Wu X-B (2014) Hybrid biogeography-based optimization for integer programming. Sci World J

30. Mirjalili S, Mirjalili SM, Lewis A (2014) Let a biogeography-based optimizer train your multi-layer perceptron. Inf Sci 269:188–209

31. Kirkpatrick S (1984) Optimization by simulated annealing: quantitative studies. J Stat Phys 34:975–986

32. Rahnamayan S, Tizhoosh HR, Salama M (2007) A novel population initialization method for accelerating evolutionary algorithms. Comput Math Appl 53:1605–1614

33. Rutkowski L, Cpalka K (2003) Flexible neuro-fuzzy systems. Neural Netw 14:554–574

34. Polat K, Sahan S, Kodaz H, Günes S (2007) Breast cancer and liver disorders classification using artificial immune recognition system (AIRS) with performance evaluation by fuzzy resource allocation mechanism. Expert Syst Appl 32:172–183

35. Mooney RJ (1996) Comparative experiments on disambiguating word senses: an illustration of the role of bias in machine learning. In: The computing research repository (CoRR), pp 82–91

36. Leon IV W (2006) Enhancing pattern classification with relational fuzzy neural networks and square BK-products. Ph.D. dissertation in computer science, Springer, FL, USA

37. Zarndt F (1995) A comprehensive case study: an examination of machine learning and connectionist algorithms. Ph.D., Department of Computer Science, Brigham Young University

38. Ene M (2008) Neural network-based approach to discriminate healthy people from those with Parkinson's disease

39. Pham HNA, Triantaphyllou E (2011) A meta-heuristic approach for improving the accuracy in some classification algorithms. Comput Oper Res 38:174–189