Comments for computers-2668380

# Reviewer 1

## 1. Comment

The value of H is only 10 in the simulation. This value may be smaller for a complex task.

## Response

Additional experiments were conducted using different values of the H paramater as well as the maximum number of generations (Nt) and two more graphs have been added at subsection 3.2 Also, the added text now reads:

*“Also, more experiments were conducted using the number of processing nodes ( H ) as well the number of generations ( N t ), to determine the stability of the proposed technique. The average classification error for the classification datasets using different values for the H parameter is shown graphically in Figure* [*fig:experimentsH*](#fig_experimentsH)*. Furthermore, the average classification error of the proposed method for different number of maximum generations N t is shown graphically in Figure* [*fig:experiments\_nt*](#fig_experiments_nt)*. The proposed technique does not show significant changes in its behavior as the critical parameters change and therefore its stability to changes is evident.”*

## 2. Comment

What type of neural network is used in the simulation. Training time comparison is useful to show the advantage of the proposed method.

## Response

Additional experiments regarding the parameters H, Nt as well the average execution time have been added at subsection 3.2 and the added text now reads:

*“Also, more experiments were conducted using the number of processing nodes ( H ) as well the number of generations ( N t ), to determine the stability of the proposed technique. The average classification error for the classification datasets using different values for the H parameter is shown graphically in Figure* [*fig:experimentsH*](#fig_experimentsH)*. Furthermore, the average classification error of the proposed method for different number of maximum generations N t is shown graphically in Figure* [*fig:experiments\_nt*](#fig_experiments_nt)*. The proposed technique does not show significant changes in its behavior as the critical parameters change and therefore its stability to changes is evident. Also, the average execution time for different values of the N t parameter is outlined in Figure* [*fig:experiments\_time*](#fig_experiments_time)*. Judging from the above experiments, it can be said that the choice of values for the critical parameters H and N t is a compromise between performance and speed for the proposed technique.”*

## 3. Comment

Can it be extended to the optimization and control issues?

## Response

The following text has been added at the end of Introduction section:

*“The proposed method utilizes a genetic algorithm in the second stage to train the neural network, but any other optimization method can be used instead of genetic algorithms. However, genetic algorithms are used because they are easily adaptable to many optimization problems, they are fault tolerant and because they can be easily parallelized using modern computing techniques.”*

## 4. Comment

There are many methods to initialize the parameters for artificial neural networks. So what is new in this paper.

## Response

The following text has been added at the end of Introduction section:

*“The proposed method seeks in its first stage to find, using an evolutionary rule technique, a range of values for the parameters of the artificial neural networks. This interval should on the one hand be small enough to speed up the method of the second stage that will be used to train the model and on the other hand the artificial neural network should have satisfactory generalization abilities within the interval of values of the first stage.”*

# Reviewer 2

## 1. Comment

The title must have a capital letter in every main word.  Bounds instead of “bounds”, Neural instead of “neural”, etc.

## Response

Corrected.

## 2. Comment

There is only one affiliation instead of three, please check and correct according to the format.

## Response

We have three affiliations.

## 3. Comment

Why was the GA selected over other techniques to train the values of the ANN? Please elaborate.

## Response

The following text has been added at the end of Introduction section:

*“The proposed method utilizes a genetic algorithm in the second stage to train the neural network, but any other optimization method can be used instead of genetic algorithms. However, genetic algorithms are used because they are easily adaptable to many optimization problems, they are fault tolerant and because they can be easily parallelized using modern computing techniques.”*

## 4. Comment

Please check Figure 1, the legend is misplaced.

## Response

Corrected.

## 5. Comment

The format of the tables does not match the MDPI format. Please check and correct.

## Response

Done.

## 6. Comment

Figures 3 and 4 do not need bold letters in the axis. Please check and correct

## Response

We have replaced the images.

# Reviewer 3

## 1. Comment

I believe that the introduction is missing a paragraph emphasizing the need to develop a new method of generating a range of values for the parameters of the artificial neural network.

## Response

The following text has been added at the end of Introduction section:

*“The proposed method seeks in its first stage to find, using an evolutionary rule technique, a range of values for the parameters of the artificial neural networks. This interval should on the one hand be small enough to speed up the method of the second stage that will be used to train the model and on the other hand the artificial neural network should have satisfactory generalization abilities within the interval of values of the first stage.”*

## 2. Comment

Due to the numerous references to literature, I think it is worth limiting the number of sources to one example of an application of a given method (e.g. instead of 'the Back Propagation method [19,20]', write 'the Back Propagation method [19]').

## Response

Done.

## 3. Comment

Write the variables in equation 7 in italics.

## Response

Done.

## 4. Comment

It is necessary to standardize the modulo notation in Table 1.

## Response

Changed to mod.

## 5. Comment

The choice of parameter values has a great impact on the results of the metaheuristics, but in the work I did not find any justification for the selected values, which are presented in Table 2. There is also no information about why you chose certain methods for comparison.

## Response

Additional experiments were conducted using different values of the H paramater as well as the maximum number of generations (Nt) and two more graphs have been added at subsection 3.2 Also, the added text now reads:

*“Also, more experiments were conducted using the number of processing nodes ( H ) as well the number of generations ( N t ), to determine the stability of the proposed technique. The average classification error for the classification datasets using different values for the H parameter is shown graphically in Figure* [*fig:experimentsH*](#fig_experimentsH)*. Furthermore, the average classification error of the proposed method for different number of maximum generations N t is shown graphically in Figure* [*fig:experiments\_nt*](#fig_experiments_nt)*. The proposed technique does not show significant changes in its behavior as the critical parameters change and therefore its stability to changes is evident.”*

## 6. Comment

I believe that in subsection 2.3 you should write at least one introductory sentence describing the pseudocode presented.

## Response

## 7. Comment

The abbreviation PSO was used and was not explained.

## Response

Done.

## 8. Comment

There is a lack of consistency in the use of proper names - sometimes you write 'Genetic Algorithms', and sometimes 'genetic algorithms'.

## Response

Corrected everywhere to genetic algorithms.

## 9. Comment

There are missing spaces before some references to literature (e.g. on line 74).

## Response

Corrected.

## 10. Comment

A dot is missing at the end of line 205.

## Response

Added.

## 11. Comment

I think the word 'Results' is missing from the name of Tables 3 and 4.

## Response

The captions have been changed to the following:

1) Experimental results for classification datasets. The values in cells indicate averare classification error as measured on test set

2) Experimental results for regression datasets. The values indicate average regression error as measured on test set.