Response to reviewer comments for the paper: INS-D-23-781

Title: Training Feedforward Neural Networks with Bayesian Hyper-Heuristics
Authors: A.N. Schreuder, A.S. Bosman, A.P. Engelbrecht, C.W. Cleghorn
Manuscript number: 1
Revision Version: 2
Editor's Decision Received Date: 7 August 2024
Revision Submission Date: 16 Aug 2024

The authors would like to thank the editor and reviewers for their constructive comments and suggestions that have helped improve the quality of this manuscript. The manuscript has undergone a final thorough revision (version 2) according to the editor's and reviewer's comments. Please see below our responses to the latest feedback. For the editor's and reviewer's convenience, we have highlighted our responses to the comments and described how each comment is addressed in the revised manuscript in blue.

Response to the Editor

Editor Comment —

- A thorough editing is required; show all changes made to the revised manuscript.
- References: max 50 entries. Make sure the references are recent and highly pertinent to the contents of the manuscript.

Reply: The authors would like to thank the editor and reviewer again for their comments and the time taken to read the manuscript. All the issues raised by the editor/reviewers are addressed in the following document. A number of grammatical and language improvements have been made, and the references are updated and now capped at 49 as requested. A summary of the changes made is given as follows:

- Changed FFNNs to FFNN in the paragraph about the BHH's applicability to different ANN types.
- Removed citation "[5]" from the sentence about FFNN topologies.
- Removed "was popularised by Werbos [4]" from the sentence about the backpropagation algorithm.
- Removed citation "[3]" from the list of references about meta-heuristics used to train FFNNs.

- Changed "dimensions" to "categories" when describing the classification scheme for hyper-heuristics. Also changed subsequent uses of "dimension" to "category".
- Removed citation "[1]" from the list of examples of hyper-heuristics in optimization.
- Changed "not individual heuristic" to "not the state of individual heuristics" when describing what the population state tracks.
- Removed citation "[2]" from the sentence about Glorot uniform sampling for weight initialization.
- Removed citations for various statistical tests: Shapiro-Wilk test, Levene's test, ANOVA test, Kruskal-Wallis test, and Tukey honest significant difference test.

References

[1] Kathryn A Dowsland, Eric Soubeiga, and Edmund Burke. A simulated annealing based hyperheuristic for determining shipper sizes for storage and transportation. *European Journal of Operational Research*, 179(3):759–774, 2007.

[2] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256, 2010.

[3] Jatinder ND Gupta and Randall S Sexton. Comparing backpropagation with a genetic algorithm for neural network training. *Omega*, 27(6):679–684, 1999.

[4] Paul John Werbos. *The roots of backpropagation: from ordered derivatives to neural networks and political forecasting*, volume 1. John Wiley & Sons, 1994.

[5] Andreas Zell. *Simulation neuronaler netze*, volume 1. Addison-Wesley Bonn, 1994.