PSO Trained Artificial Neural Networks

1st Anwar R Rizvi
Post Graduate Student; ID: H00315584
School of Mathematical and Computer Sciences
arr8@hw.ac.uk

2nd Jyotish Joseph Thomas
Post Graduate Student; ID: H0315392
School of Mathematical and Computer Sciences
jjt6@hw.ac.uk

Abstract - Architectural design and weights optimization of Artificial Neural Networks is a challenging task in machine learning issues. In this study, we explore usage of Particle Swarm Optimization algorithm to optimize neural network architecture and weights for enhancing accuracy and efficiency of the network.

I. Introduction

Artificial Neural Networks (ANN) are adaptive and flexible and learn from their surroundings adapting to internal and external parameters. It can be used to create model to solve complex computational problems from large data set with multiple correlated variables. ANN's capability is heavily dependent on ANN parameters of activation function, number of network layers and neurons in the layers.

Particle Swarm Optimization (PSO) is a popular stochastic global optimization technique inspired from ways in which biological organisms forage for food. It is originally attributed to Kennedy, Eberhart and Shi. It represents a population-based adaptive optimization technique that is influenced by its hyper parameters (eg. Swarm size, inertia, cognitive and social weights). It has been proved efficient in solving global optimization and engineering problems. PSO introduces a population of particles representing solution space and these particles are moved in space by mathematical formulae until they converge to an appropriate solution. Every particle has N-dimensional position and velocity vectors. Using a particle's position vector, a fitness function gives the fitness of the particle. The objective is to maximize the fitness of the particles.

In this article, we present a method for parameter optimization based on PSO and its application to Artificial Neural Network (ANN) training so as to overcome main two defects of traditional ANN, which are slowness due to back-propagation using derivative functions and high risk of local minima stagnation. The PSO based ANN network enhances the accuracy of output prediction in real life classification problems like image classifications, weed detection, fruit ripeness detection etc. We assessed the performance of PSO trained ANN on a set of six artificial fitness functions. The results obtained were compared and discussed with RMSE evaluation metric.

Training neural network using PSO provides less training time and high accuracy. The proposed technique is evaluated with six data sets covering six different functions being optimized using PSO and ANN. The results show that the technique can predict outputs with RMSE in the range of 0-2% using ANN trained using PSO.

Main benefits of PSO includes easy implementation, robustness to change in parameters and efficiency compared to other heuristic algorithms such as genetic algorithm. PSO can be applied to many nondifferentiable, non-linear problems to get quick and efficient results.

The article is organized as follows: Section 2 presents the proposed design and methodology is described in Section 3. Section 4 presents analyzes and results obtained from our investigations. Finally, in Section 5 we summarize our conclusions.

II. PROGRAM DEVELOPMENT RATIONALE

III. METHODS

ANN Architecture: Artificial Neural Network architecture was built based on standard framework as described in Fig 2 of VII. Annex I - Tables and Figures.

PSO Algorithm: The PSO algorithm for the global best network topology is implemented as shown in Fig 3 of VII. Annex I - Tables and Figures.

The algorithm is clear indicator of the competitive nature of the particles in space. In addition, global best swarm converges faster due to the social nature of the velocity update.

IV. RESULTS

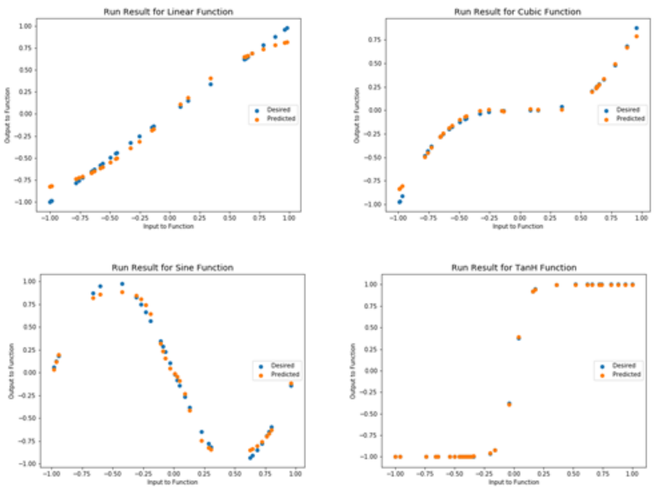
In order to select the appropriate ANN and PSO hyper parameters and to evaluate the performance of the algorithms, we initially executed the below listed investigation runs using provided datasets. The results of the experiments are provided in Table 1 of VII. Annex I - Tables and Figures. In Figures 1-6, we present boxplots of the resulting RMSE's for the algorithms applied on the respective functions data sets.

The results of the runs with optimized hyper parameters are provided in Table 2 – Investigation summary with selected PSO and ANN hyper parameter runs

Investigations were carried out for each of the given functions i.e. Linear, Cubic, Sine, TanH, XOR, Complex.

1. Varying Swarm Size - [5, 10, 15, 30]
2. Varying Inertia weight (Alpha) - [0.2, 0.6, 0.9, 1.2, 1.4, 1.6]
3. Varying Cognitive weight (Beta) - [0.4, 0.8, 1.2, 1.6, 2.0, 2.4]
4. Varying Social weight (Delta) - [0.4, 0.8, 1.2, 1.6, 2.0, 2.4]
5. Varying Activation Function - ['Null', 'Sigmoid', 'Hyperbolic Tangent', 'Cosine', 'Gaussian']
6. Varying # Layers in ANN - [7],[7,7],[7,7,7],[7,7,7,7]]
7. Varying # Neurons in ANN – [[1,1],[5,3],[8,5],[10,8]]

Fig 1 – Predicted vs. Desired Outputs



V. DISCUSSION AND CONCLUSION

All optimization problems have its own features which should be considered while optimizing the problem. In past few years a lot of optimization techniques have been developed. However, any one algorithm so far can not successfully optimize all optimization problems (Wolpert et al. 1997). Particle swarm optimization inspired from social behavior of biological swarms was evaluated based on six fitness functions. PSO proved to be an optimal solution in prediction of outputs. The prediction accuracy can be improved by more training. However, achieving further accuracy needs more research.

This work is another application of PSO. Primarily focused on identifying the significance of usage of PSO to train Artificial Neural Networks. However, further research can be undertaken to research in depth on the impact of different variants of PSO by varying the neighborhood.

VI. References

[1] J. Kennedy, R.C. Eberhart, et al., “Particle swarm optimization”, In Proceedings of IEEE international conference on neural networks, volume 4, pages 1942–1948. Perth, Australia, 1995.

[2] Y. Shi and R. Eberhart., “A modified particle swarm optimizer”, In Evolutionary Computation Proceedings, 1998. IEEE World Congress on Computational Intelligence., The 1998 IEEE International Conference on, pages 69–73. IEEE, 2002.

VII. Annex I - Tables and Figures

Fig 2 – Flowchart of Artificial Neural Network Architecture

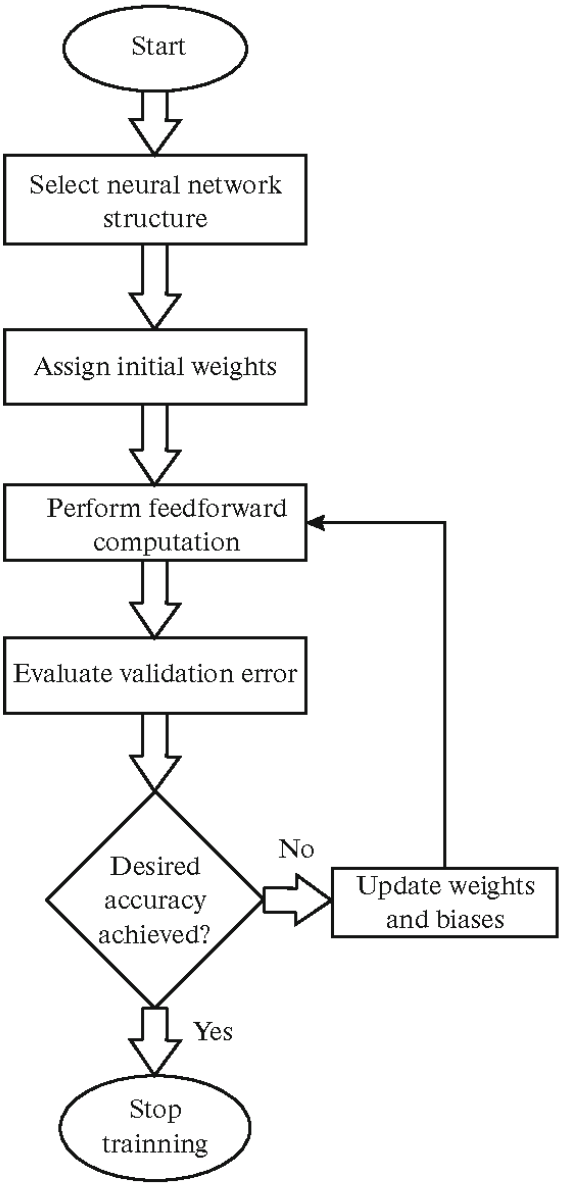


Fig 3 – Flow chart of Particle Swarm Optimization Algorithm

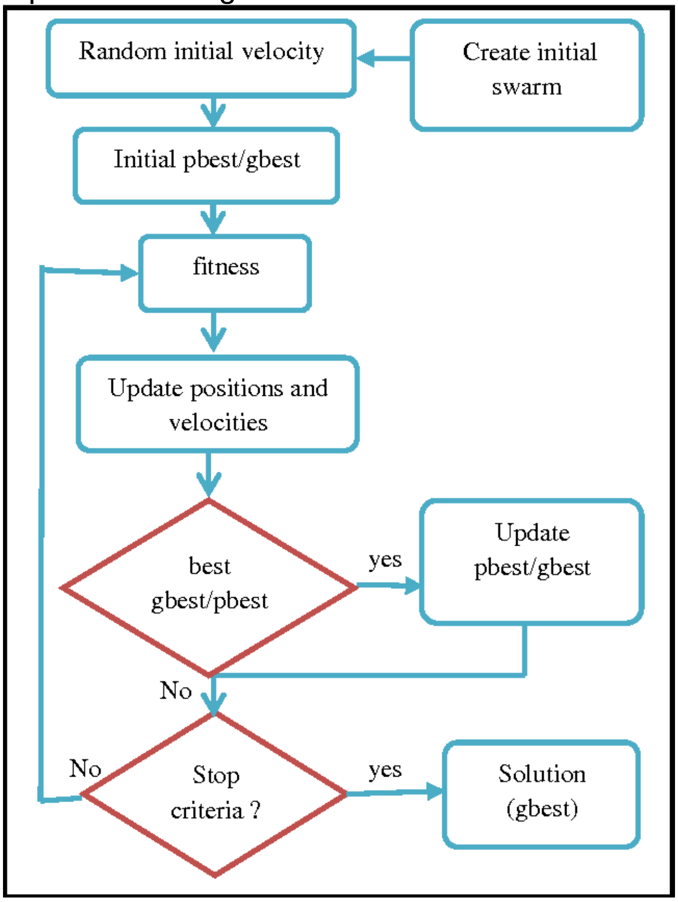


Fig 4 – Charts for Activation Functions

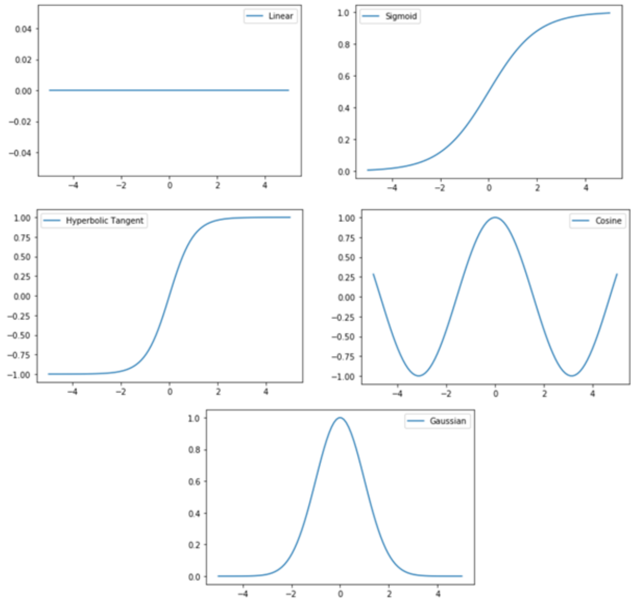


Table 1 – Investigation summary for varying PSO and ANN hyper parameter runs

Fig 1. Graphs for Linear Function

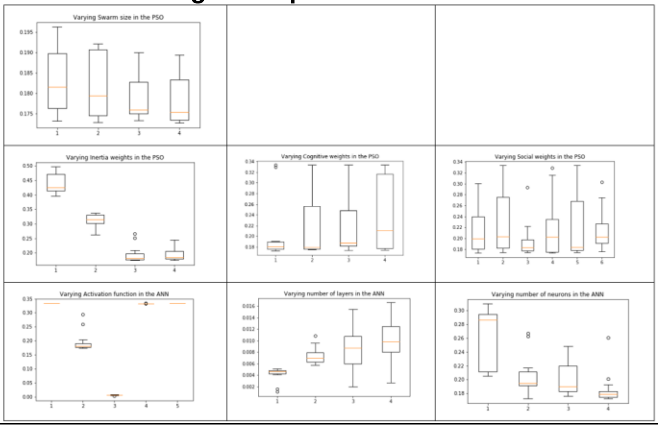


Fig 2. Graphs for Cubic Function

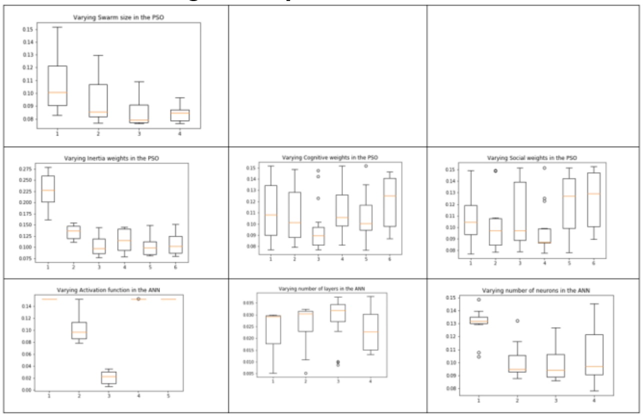


Fig 3. Graphs for Sine Function

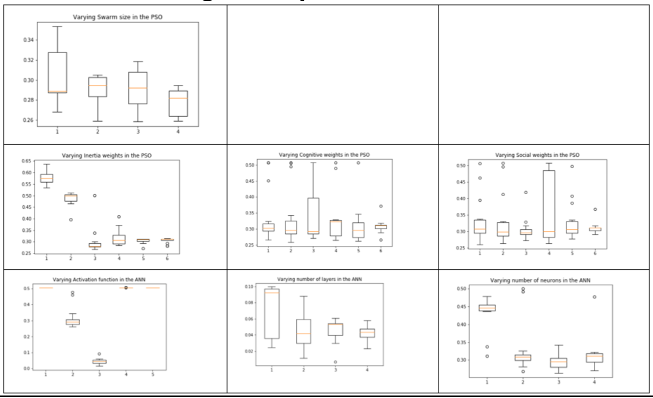


Fig 4. Graphs for Tanh

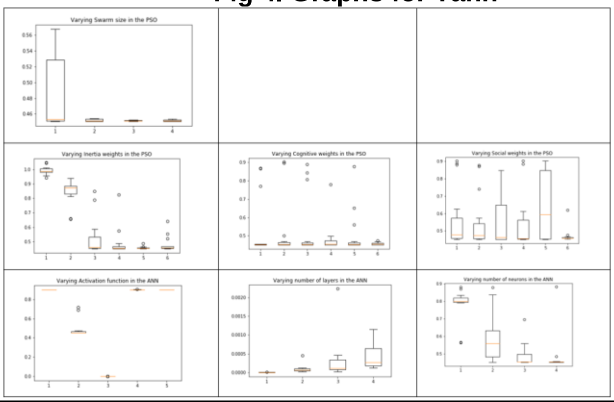


Fig 5. Graphs for XOR Function

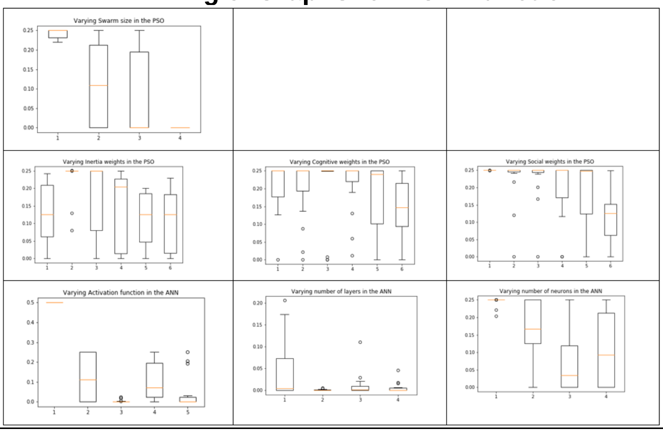


Fig 6. Graphs for Complex Function

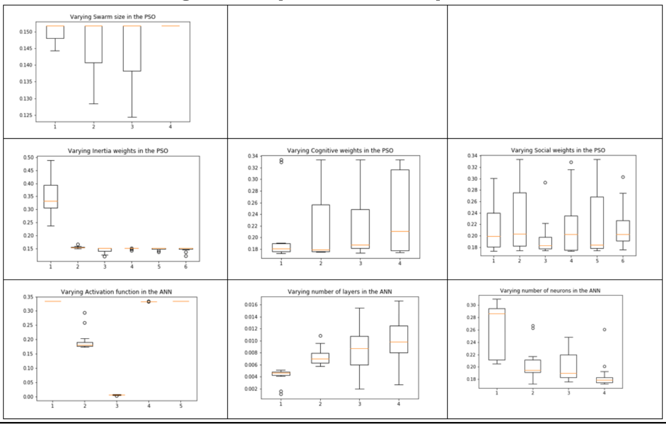


Table 2 – Investigation summary with selected PSO and ANN hyper parameter runs