

Function Approximations with Multilayer Perceptrons and Simultaneous Recurrent Networks

^{1,2}Ganesh K. Venayagamoorthy, ³Xiaohua Wang, ^{1,3}Manash Buragohain, ³Amit Gupta

¹Real-Time Power and Intelligent Systems Laboratory, ²Department of Electrical and Computer Engineering, ³Department of Mechanical and Aerospace Engineering
University of Missouri-Rolla, MO 65409, USA
gkumar@ieee.org

Keywords: Simultaneous Recurrent Networks, Multilayer Perceptrons, Particle Swarm Optimization (PSO), Backpropagation Training, Function Approximation

Abstract

Approximation of highly nonlinear functions is an important area of computational intelligence. The relevance of this area can be found traced to numerous fields, from controlling nonlinear complex dynamics, to many practical and difficult applications such as a brain-like intelligent control and planning. It has been widely accepted for a considerable period that multilayer feedforward neural networks or Multilayer Perceptrons (MLPs) with sigmoidal activation functions have been termed as "universal function approximators" [1]. The inability to approximate non-smooth functions by using conventional MLPs and attempts to replicate the human brain resulted in many new types of ANN such as the simultaneous recurrent networks (SRNs) [2]. The focus of this paper is threefold, namely: First, to demonstrate that a neural network can learn or approximate another neural network. Second, to compare the capabilities of a Multilayer Perceptron with a Simultaneous Recurrent Network for approximating or learning any randomly generated non-linear function. Third, to demonstrate and utilize other computational intelligence paradigms such as particle swarm optimization (PSO) to obtain better convergence in training of MLP and SRN neural networks.

A non-linear function f_1 is generated using a MLP. This function f_1 is approximated by both a MLP trained with conventional Backpropagation (BP) and a SRN trained with PSO [3]. The nonlinearity and complexity of f_1 is varied by changing the size of the dataset generated. Similarly, another non-linear function f_2 is generated using a SRN. This function f_2 is approximated by both a MLP trained with BP (MLPBP) and a SRN trained with PSO (SRNPSO). The comparison of the MLPBP and SRNPSO is based on two factors, the number of epochs taken and the average mean square error achieved. Figures 1 and 2 show the structures of the MLP and SRN respectively.

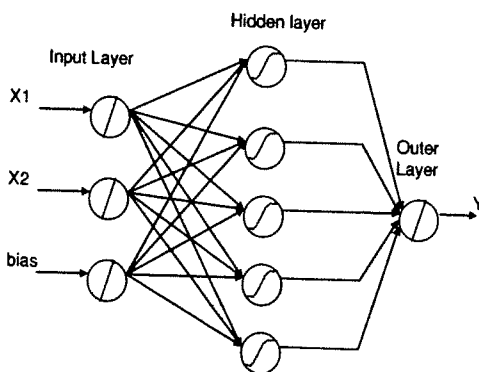


Fig 1. The structure of MLP

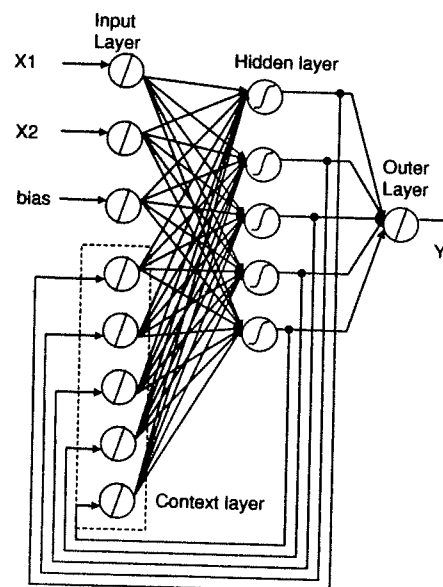


Fig. 2. The structure of SRN

In addition, the MLP is also trained with PSO (MLPPSO) and compared to SRNPSO. Different datasets sizes (20, 50 and 100) are generated by the MLP and SRN functions f_1 and f_2 respectively. The average mean square error (MSE) during testing and the number of average epochs required by MLPBP, MLPPSO and SRNPSO are given over 100 trials in Tables 1 and 2 respectively. The use of MLP and SRN to predict the Mackey-Glass time series has also been carried out and the results show that the SRN predicts better compared to the MLP. Figures 3 and 4 show the results of the MLPPSO and SRNPSO respectively.

From the results obtained in this study, it can be inferred that generally the SRN achieves higher success rate than the MLP. The SRN outperforms both MLP generated and SRN generated non-linear functions. Further, as the complexity of the function is increased, the SRN is seen to be more robust compared to the MLP. The other revealing fact of the analysis is that, in most of the cases, the rate of success of MLP when approximating MLP generated function is far lesser than that for SRN generated function. However, the effect of the function generator is lesser on the SRN approximations. So, it can be concluded that wherever a MLP carries out a good approximation, a SRN will also carry out a good approximation, but the vice-versa is not true.

Table 1: Average MSE over 500 epochs

	MLP generator MSE ($\times 10^{-5}$)	SRN generator MSE ($\times 10^{-5}$)
MLPBP 20 data points	7.77	12.93
MLPPSO 20 data points	1.51	2.37
SRNPSO 20 data points	2.15	4.75
MLPBP 50 data points	10.60	13.91
MLPPSO 50 data points	2.83	3.01
SRNPSO 50 data points	4.11	3.79
MLPBP 100 data points	6.39	12.98
MLPPSO 100 data points	1.76	3.10
SRNPSO 100 data points	3.04	3.64

Table 2: Average epochs over to obtain MSE of 5×10^{-4}

	MLPBP	MLPPSO	SRNPSO
MLP generator 20 data points	94	30	34
SRN generator 20 data points	125	38	48
MLP generator 50 data points	65	31	41
SRN generator 50 data points	96	36	42
MLP generator 100 data points	34	33	33
SRN generator 100 data points	78	35	39

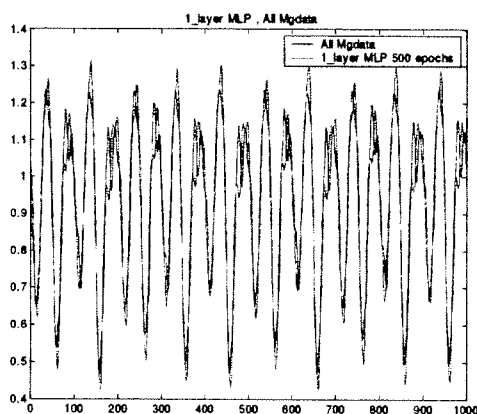
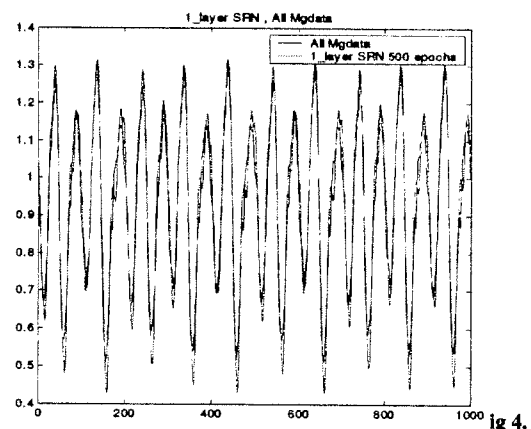


Fig 3. Mackey-Glass time series prediction by MLPPSO



ig 4. Mackey-Glass time series prediction by SRNPSO

References

- [1] Hines JW, "A Logarithmic Neural Network Architecture for Unbounded on- Linear Function Approximation" *International Conference on Neural Networks*, 1996. Washington, DC, USA. Pages: 174.
- [2] Pang X, Werbos P, "Neural Network Design for J Function Approximation in Dynamic Programming", *Math. Modelling and Scientific Computing*, Vol. 5, No. 2/3, 1996.
- [3] Kennedy J, Eberhart R, "Empirical study of particle swarm optimization", *Proceedings of the 1999 Congress on Evolutionary Computation*, vol. 3, Pages: 1950.