

On Generalization Ability of Cascaded Neural Net Architecture

Joarder KAMRUZZAMAN[†] Yukio KUMAGAI^{††} Hiromitsu HIKITA^{†††}

[†]Dept. of Electrical & Electronic Engg.
Bangladesh Univ. of Engg. & Tech.
Dhaka 1000, Bangladesh

^{††}Dept. of Comp. Science & Syst. Engg.
^{†††}Dept. of Mechanical Syst. Engg.
Muroran Inst. of Tech.
27-1 Mizumoto-cho, Muroran 050, Japan

Abstract - We present construction of a 3-layer feedforward network called cascaded network and make a performance comparison with the conventional Backpropagation network [1] in terms of generalization ability of the trained networks. Cascaded networks are trained to realize category classification employing binary input vectors and locally represented target output vectors. Empirical study in [2] shows that when a neural network is trained to classify binary input vectors into locally represented binary target output vectors, better saturation of hidden outputs in response to the training set yields better robustness of the network, and in conventional Backpropagation network hidden outputs usually do not get saturated in response to the training set. Based on this observation a 3-layer cascaded network is constructed by cascading two 2-layer networks trained independently by delta rule [3] whose intermediate layer can be viewed as hidden layer and is trained to attain preassigned saturated outputs in response to the training set. Each input-output pair is associated with a preassigned code. Each 2-layer network can thus be considered as a separate module; the first module is trained with the training pattern as the binary input signal and the corresponding preassigned code as the teacher signal. The second module is trained with this preassigned code as the input signal and the desired binary output of the corresponding training pattern as the teacher signal. Thus the first module maps the training patterns onto the preassigned codes and the second maps these preassigned codes onto the desired outputs. After cascading, outputs of the first module become inputs to the second module and thus the intermediate layer of the resultant 3-layer cascaded network can be considered as the hidden layer which is trained to attain predefined saturated outputs as the internal representation of the training set. For linearly separable tasks cascaded network can be built straightforwardly as described earlier, and for nonlinearly separable task, cascaded network is constructed by employing high order cross product inputs at the input layer. These high order cross product inputs are implemented as the mutually disjoint cross products of the components of the binary input vectors received from the training patterns. Experimenting with character recognition problems we demonstrate by simulation results that, for both linearly and nonlinearly separable tasks with binary input and locally represented binary target output vectors, cascaded network yields generalization ability far better than that of Backpropagation network. The better performance of cascaded network is duo to the attainment of saturated hidden outputs in response to the training set.

References

- [1] D. E. Rumelhart, J. L. McClelland and the PDP research Group, *Parallel Distributed Processing* vol. 1, M.I.T. Press, 1986.
- [2] M. Hamamoto, J. Kamruzzaman, Y. Kumagai and H. Hikita, "Generalization ability of feedforward neural network trained by Fahlman and Lebiere's learning algorithm," *IEICE Trans. Fundamentals*, Japan, vol. E75-A, no. 11, pp. 1597-1601, Nov. 1992.
- [3] B. Widrow and M. E. Hoff, "Adaptive switching circuits," *IRE WESTCON Convention Record*, part 4, pp. 96-104, 1960.