A Selective Learning Alogrithm for Nonlinear Synapses in Multilayer Neural Networks

by Minoru Fusakawa

(Supervised by Professor Kenji Nakayama and Akihiro Hirano)

ABSTRACT

Network size of neural networks is highly dependent on activation functions. A trainable activation function has been proposed, which is a linear combination of some basic functions. The activation functions and the connection weights are simultaneously trained. An 8 bit parity problem can be solved by using a single output unit and no hidden unit. This model can be applied to multilayer neural networks. Furthermore, nonlinear functions are introduced for the unit input in order to expand transfer function representation. The activation functions and the nonlinear functions can be simultaneously trained. More complex pattern classification and function approximation problems can be solved with a small number of units and fast convergence.

When the nonlinear functions are realized using a high-order polynomial, however, a huge number of terms are required for the high-dimensional input. This causes large size and complex networks and slow convergence. These situation will occur in actual applications. Therefore, it is very important to select the minimum useful terms in a learning process.

In this paper, we propose a selecting method of the useful terms based on the genetic algorithm (GA) modified by the internal information of the network. A chromosome is assigned to a polynomial in the networks. A gene is assigned to each term of the polynomial. When the gene is 1, the corresponding term is used, and the gene is 0, it is not used. A concrete selecting process is as follows: Step 1: K individuals, that is multilayer neural networks, are generated using connection weights and location of the gene, having 1, which are randomly determined. The number of the genes having 1 is limited to M. Step 2: The K individuals generated in Steps 1 and 3 are trained by the simultaneous learning algorithm for the connection weights and the activation functions. The output error is evaluated. If the error satisfies the requirement and the number of the genes having 1 is well reduced, then the learning is completed. Otherwise, the process moves to the next step. Step 3: K' good individuals evaluated by the output error are selected. Two of them are combined to generate individuals in the next generation. Let G_1 and G_2 be the good individuals to be combined. Suppose G_1 is better than G_2, the gene with the maximum connection weight in G_1 is selected to be used in the new individual. Next, the gene is selected in the same way from G_2. This selection is repeated alternately M times. Thus, the maximum number of the genes having 1 is M. In this way, K new individuals are generated.

The proposed method was applied to many problems, including 2-dimensional function approximation, 2-dimensional pattern classification and 3-dimensional pattern classification. In any cases, the proposed method can select the suitable terms for each problem. The number of the selected terms is well reduced and convergence is stable compared with the method using the ordinary genetic algorithm.