CUMULATIVE NEGATIVE TRANSFER DURING SUCCESSIVE TRAINING: ANALYSIS OF A SECOND SEQUENTIAL LEARNING PROBLEM

Adel M. Abunawass and William S. Maki
Departments of Computer Science and Psychology
North Dakota State University
Fargo, North Dakota 58105-5075

Abstract. Adaptive networks can be easily trained to associate arbitrary input and output patterns. When subgroups of patterns (lists) are presented sequentially, however, a network tends to "unlearn" previously acquired associations while learning new associations. A second form of sequential learning problem is reported in this paper. Learning of each successive list of pattern pairs becomes progressively more difficult. Evidence for this cumulative negative transfer was obtained from simulations using back-propagation of errors to train multilayer networks. The cause of the problem appears to be the development of extreme weights during learning of new lists. Unbounded weights may be a liability for the back-propagation algorithm.

DYNAMIC NODE CREATION IN BACKPROPAGATION NETWORKS

Timur Ash

Department of Computer Science and Engineering
University of California, San Diego
La Jolla, CA 92093
ash%cs.ucsd.edu

Abstract

Large backpropagation (BP) networks are very difficult to train. This fact complicates the process of iteratively testing different sized networks (i.e., networks with different numbers of hidden layer units) to find one that provides a good mapping approximation. This paper introduces a new method called Dynamic Node Creation (DNC) that attacks both of these issues (training large networks and testing networks with different numbers of hidden layer units). DNC sequentially adds nodes one at a time to the hidden layer(s) of the network until the desired approximation accuracy is achieved. Simulation results for parity, symmetry, binary addition, and the encoder problem are presented. The procedure was capable of finding known minimal topologies in many cases, and was always within three nodes of the minimum. Computational expense for finding the solutions was comparable to training normal BP networks with the same final topologies. Starting out with fewer nodes than needed to solve the problem actually seems to help find a solution. The method yielded a solution for every problem tried. BP applied to the same large networks with randomized initial weights was unable, after repeated attempts, to replicate some minimum solutions found by DNC.

TECHNIQUES FOR SYNTHESIZING PIECEWISE LINEAR AND QUADRATIC NEURAL NETWORK CLASSIFIERS

Joseph B. Attili

PAR Government Systems Corporation
220 Seneca Turnpike
New Hartford, New York 13413

Abstract

Neural network (NN) classifiers have been applied to numerous practical problems of interest. A very common type of NN classifier is the multi-layer perceptron, trained with back propagation. Although this learning procedure has been used successfully in many applications, it has several drawbacks including susceptibility to local minima and excessive convergence times.

This paper presents two alternatives to back propagation for synthesizing NN classifiers. Both procedures generate appropriate network structures and weights in a fast and efficient manner without any gradient descent. The resulting decision rules are optimal under certain conditions; the weights obtained via these procedures can be used "as is" or as a starting point for back propagation.

Fast Learning Process of Multi-Layer Neural Nets Using Recursive Least Squares Technique

Mahmood R. Azimi-Sadjadi and Stuart Citrin

Department of Electrical Engineering
Colorado State University
Fort Collins, CO 80523
Tel: (303)-491-7956

Abstract:

In this paper a new approach for learning process of multi-layer perceptron neural networks using Recursive Least Squares (RLS) technique is proposed. This method minimizes the sum of the square of the errors between the actual and the desired output values recursively. The weights in the net are updated upon the arrival of a new training sample and by solving a system of normal equations using matrix inversion lemma. To determine the desired target in the hidden layers an analog of back propagation strategy used in the conventional learning algorithms is developed. This permits the application of the learning procedure to all the other layers. Simulation results on the Exclusive-OR example are obtained which indicate significant (an order of magnitude) reduction in the total number of iterations when compared with those of conventional techniques.