LETTER Communicated by Scott Fahlman

Efficient Adaptive Learning for Classification Tasks with Binary Units

J. Manuel Torres Moreno
Mirta B. Gordon
Département de Recherche Fondamentale sur la Matière Condensée, CEA Grenoble, 38054 Grenoble Cedex 9, France

This article presents a new incremental learning algorithm for classification tasks, called NetLines, which is well adapted for both binary and real-valued input patterns. It generates small, compact feedforward neural networks with one hidden layer of binary units and binary output units. A convergence theorem ensures that solutions with a finite number of hidden units exist for both binary and real-valued input patterns. An implementation for problems with more than two classes, valid for any binary classifier, is proposed. The generalization error and the size of the resulting networks are compared to the best published results on well-known classification benchmarks. Early stopping is shown to decrease overfitting, without improving the generalization performance.

1 Introduction

Feedforward neural networks have been successfully applied to the problem of learning pattern classification from examples. The relationship of the number of weights to the learning capacity and the network's generalization ability is well understood only for the simple perceptron, a single binary unit whose output is a sigmoidal function of the weighted sum of its inputs. In this case, efficient learning algorithms based on theoretical results allow the determination of the optimal weights. However, simple perceptrons can generalize only those (very few) problems in which the input patterns are linearly separable (LS). In many actual classification tasks, multilayered perceptrons with hidden units are needed. However, neither the architecture (number of units, number of layers) nor the functions that hidden units have to learn are known a priori, and the theoretical understanding of these networks is not enough to provide useful hints.

Although pattern classification is an intrinsically discrete task, it may be cast as a problem of function approximation or regression by assigning real values to the targets. This is the approach used by backpropagation and

