

Scaling and Generalization in Neural Networks: A Case Study

Authors: Subutai Ahmad, Gerald Tesauro

Abstract: The issues of scaling and generalization have emerged as key issues in current studies of supervised learning from examples in neural networks. Questions such as how many training patterns and training cycles are needed for a problem of a given size and difficulty, how to represent the input and how to choose useful training exemplars, are of considerable theoretical and practical importance. Several intuitive rules of thumb have been obtained from empirical studies, but as yet there are few rigorous results. In this paper we summarize a study of generalization in the simplest possible case-perceptron networks learning linearly separable functions. The task chosen was the majority function (i.e. return a 1 if a majority of the input units are on), a predicate with a number of useful properties. We find that many aspects of generalization in multilayer networks learning large, difficult tasks are reproduced in this simple domain, in which concrete numerical results and even some analytic understanding can be achieved.