Prediction Intervals for Neural Networks via Nonlinear Regression

Richard De Veaux, Jennifer Schumi, Jason Schweinsberg, David Shellington
Department of Mathematics, Williams College
and
Lyle H. Ungar
Department of Computer and Information Science, University of Pennsylvania

Abstract

Standard methods for computing prediction intervals in nonlinear regression can be effectively applied to neural networks when the number of training points is large. However, simulations show that thses methods can generate unreliable prediction intervals on smaller data sets because when the network is trained to convergence. Stopping the training algorithm prior to convergence, to avoid over-fitting, reduces the effective number of parameters effectively, but can lead to prediction intervals that are too wide. We present an alternative approach to estimating prediction intervals which uses weight decay to fit the network and show that this method is effective on a wide range of problems.

KEY WORDS: Nonparametric regression, smoothing, high-dimensional data, backpropagation.

1 Introduction

Multilayer feedforward neural networks are flexible models that are widely used to model high-dimensional, nonlinear data. The models typically contain many parameters, sometimes as many or more parameters as observations. Because of this, the networks are uninterpretable, and are often regarded as "black box" models. Unfortunately, many users, in spite of, (or because of) this, take the predictions given by the network on faith, without regard to the uncertainly inherent in the predictions. Few users of neural networks compute prediction intervals even though,