Variable Architecture Bayesian Neural Networks

Silvia Bozza
Department of Statistics
University of Venice
silvia.bozza@unive.it

Pietro Mantovan
Departiment of Statistics
University of Venice
pietro.mantovan@unive.it

Rosa Anna Schiavo
Departiment of Statistics
University of Venice
rschiavo@vivaldi.dst.unive.it

Keywords: feed-forward neural networks, genetic algorithms, mcmc, evolutionary monte carlo algorithm.

Abstract

A crucial problem which arises when dealing with Bayesian neural networks is that of determining their most appropriate size, expressed in terms of number of computational units and/or connections. In fact, too small a network may not be able to learn the sample data, whereas one that is too large may give rise to overfitting phenomena and cause poor "generalization" performance. A few solutions have been proposed in the literature to solve this problem, such as the use of a geometric prior probability on the number of hidden units (Müller and Rios Insua, 1998), thereby favouring smaller-size networks, and a reversible jump algorithm to move between architectures having a different number of hidden units (Rios Insua and Müller, 1998).

In this work we propose a variable architecture model where input-to-hidden connections and, therefore, hidden units are selected by using a variant of the Evolutionary Monte Carlo (EMC) algorithm developed by Liang and Wong (2000). The EMC algorithm works by simulating a population of models in parallel. The population is updated by three genetic operators: selection, crossover and mutation. For Bayesian learning of parameters we implement a hybrid Markov chain Monte Carlo algorithm which consists in a Gibbs-sampling algorithm with Metropolis steps to update the network parameters and an Evolutionary Monte Carlo step to select the architecture (Bozza, Mantovan, Schiavo, 2002). The most attractive features of genetic algorithms and simulated annealing are thus incorporated into the framework of Markov chain Monte Carlo.

Some experimental results which demonstrate the effectiveness of the proposed approach are reported.

References

Bozza, S., Mantovan, P. and Schiavo, R.A. (2002): Evolutionary Model Selection in Bayesian Neural Networks. Proceedings of 26th Annual Conference of the German Classification Society, Springer.

Liang, F. and Wong, W.H. (2000): Evolutionary Monte Carlo: Applications to C_p Model Sampling and Change Point Problem. *Statistica Sinica ,10, 317-342.*

Müller, P. and Rios Insua, D. (1998): Issues in Bayesian Analysis of Neural Network Models. *Neural Computation, 10, 749–770.*

Rios Insua, D. and Müller, P. (1998): Feedforward Neural Networks for Nonparametric Regression. In Dey D.K., Müller, P. and Sinha D. (Eds.) *Practical Nonparametric and Semiparametric Bayesian Statistics.* Springer, 181–194.