Time Series Prediction by Neural Networks

Frank M. Thiesing, Oliver Vornberger
University of Osnabrück

Keywords: neural networks, time series prediction, back-propagation

Introduction

Time series prediction for economic processes is a topic of increasing interest. In order to reduce stock-keeping costs, a proper forecast of the demand in the future is necessary. We use artificial neural networks for a short term forecast for the sale of articles in supermarkets. The nets are trained on the known sales volume of the past for an entire group of related products. Additional information like changing prices is also given to the net to improve the prediction quality. The net is trained on a window of inputs describing a fixed set of recent past states by the *back-propagation* algorithm.

Artificial neural networks

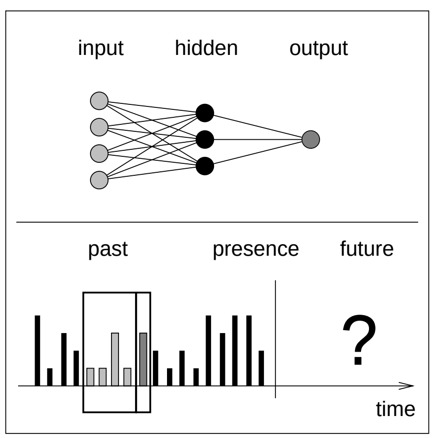

Figure 1: Feedforward multilayer perceptron for time series prediction

Artificial neural networks consist of simple calculation elements, called neurons, and weighted connections between them. In a *feedforward multilayer perceptron* (figure 1) the neurons are distributed in layers and a neuron from one layer is fully connected only to each neuron of the next layer. Values are given to the neurons in the input

layer; the results are taken from the output layer. The outputs of the input neurons are propagated through the hidden layers of the net.
Such a feedforward multilayer perceptron can approximate any function after a suitable amount of training. For that discrete values of this function are presented to the net. The net is expected to learn the function rule. The behaviour of the net is changed by modification of the weights and bias values.

Sale forecast by neural networks

In our project we use the sale information of 53 articles of the same product group in a supermarket. The information about the number of sold articles and the sales revenues in DM are given weekly starting September 1994. In addition there are advertising campaigns for articles often combined with temporary price reductions. Such a campaign lasts about two weeks and has a significant influence on the demand on this article. The sale and average price for one article are shown in figure 2.

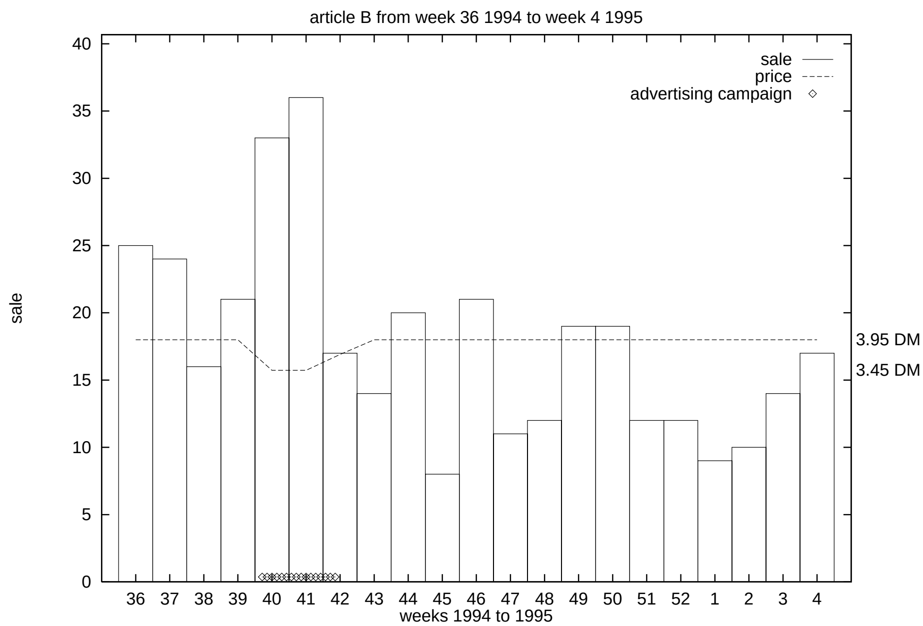

Figure 2: sale of article with advertising

The aim is to forecast the sale of an article for the next week by neural networks. For prediction the past information of n recent weeks is given to the input layer. The only result in the output layer is the sale for the next week. So there is a window of n weeks in the past and one in the future. Both the input and output together are called a training pair. One training of all training pairs is called an *epoch*.
Because all the considered articles belong to one product group, we have a quite constant sales volume of all products. An increasing sale of one article leads to a decrease of the

other products. Because of this reason, we train one neural net for the prediction of each article with the information of all articles in the input layer. We have the three items sale, advertising and price for three weeks in the recent past for 53 articles. So our nets have nearly 600 input neurons and between 50 to 100 neurons within the hidden layer. This leads to enormous training times.

Parallelization

For enhancement the back-propagation algorithm has been parallelized in different manners: First the training set can be partitioned for the batch learning implementation. The neural network is duplicated on every processor of the parallel machine, and each processor works with a subset of the training set. After each epoch of training the training results are broadcasted and merged.
A second approach is the parallel calculation of the matrix products that are used in the learning algorithm. The neurons on each layer are partitioned into p disjoint sets and each set is mapped on one of the p processors. The new activations are distributed after each training pair. We have implemented this on-line training in two variants: For the first parallelization one matrix product is not determined on one processor, but it is calculated while the subsums are sent around on the processor cycle. The second method tries to reduce communication. Therefor it leads to an overhead in both storage and number of computational operations.
The parallel implementations take place on parallel PARSYTEC systems: a T800 Multicluster and a PowerXplorer. The parallelizations run both with the runtime envirionments PARIX and PVM/PARIX.

References

[1] Ulrich Middelberg, Frank M. Thiesing, Oliver Vornberger. *Parallele Backpropagation.* in R. Flieger, R. Grebe (eds.) *Parallele Datenverarbeitung aktuell: TAT '94*, pp. 419-427, IOS Press 1994.

[2] Frank M. Thiesing, Ulrich Middelberg, Oliver Vornberger. *Parallel Back-Propagation for the Prediction of Time Series.* 1st PVM Users' Group Meeting, Rome, October 9 – 11, 1994. http://www.netlib.org/pvm3/epvmug94/contrib/thiesing_a.ps

Contact address

Contact:	Dipl.-Math. Frank M. Thiesing, Prof. Dr. Oliver Vornberger
Institution:	University of Osnabrück
	Department of Mathematics and Computer Science
Mail:	D-49069 Osnabrück
	Germany
Phone:	+49 541 969 2558
Fax:	+49 541 969 2770
Email:	frank@informatik.uni-osnabrueck.de
WWW:	http://brahms.informatik.uni-osnabrueck.de/prakt/prakt.html