

# Sales Forecasting Using Neural Networks

Frank M. Thiesing and Oliver Vornberger  
Department of Mathematics and Computer Science  
University of Osnabrück  
D-49069 Osnabrück, Germany  
frank@informatik.uni-osnabrueck.de

## Abstract

*Neural networks trained with the back-propagation algorithm are applied to predict the future values of time series that consist of the weekly demand on items in a supermarket. The influencing indicators of prices, advertising campaigns and holidays are taken into consideration. The design and implementation of a neural network forecasting system is described that has been developed as a prototype for the headquarters of a German supermarket company to support the management in the process of determining the expected sale figures. The performance of the networks is evaluated by comparing them to two prediction techniques used in the supermarket now. The comparison shows that neural nets outperform the conventional techniques with regard to the prediction quality.*

## 1. Introduction

A central problem in science is predicting the future of temporal sequences. Examples range from forecasting the weather to anticipating currency exchange rates. The desire to know the future is often the driving force behind the search for laws in science and economics.

In recent years many sophisticated statistical methods have been developed and applied to forecasting problems [8], however, there are two major drawbacks to these methods. First for each problem an individual statistical model has to be chosen that makes some assumptions about underlying trends. Second the power of deterministic data analysis can be exploited for single time series with some hidden regularity (though strange and hard to see but existent), however, this approach fails for multidimensional time series with mutual non-linear dependencies.

As an answer to the weakness of statistical methods in forecasting multidimensional time series an alternative approach gains increasing attraction: neural networks [5]. The practicability of using neural networks for economic forecasting has already been demonstrated in a variety of applications, such as stock market and currency exchange rate prediction, market analysis and forecasting time series of political economy [6, 3, 1, 2].

The approaches are based on the idea of training a feed-forward multi-layer network by a supervised training algorithm in order to generalize the mapping between the input and output data and to discover the implicit rules governing the movement of the time series and predict its continuation in the future. Most of the proposals deal with one or only few time series.

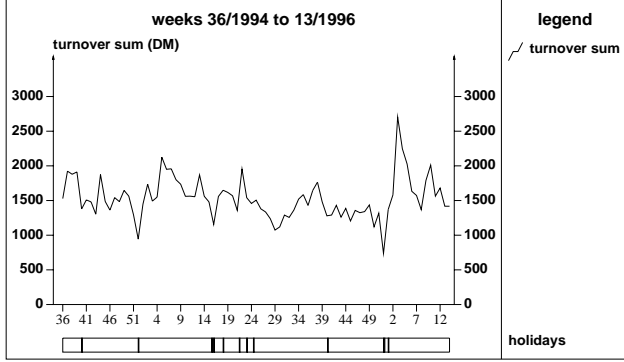
In this paper, neural networks trained with the *back-propagation* algorithm [7] are applied to predict the future values of 20 time series that consist of the weekly demand on items in a German supermarket. An appropriate network architecture will be presented for a mixture of both explanatory and time series forecasting. Unlike many other neural prediction approaches described in the literature, we compare the forecasting quality of the neural network to two prediction techniques currently used in the supermarket. This comparison shows that our approach produces good results.

## 2. Time Series Considered

The time series used in this paper consist of the sales information of 20 items in a product group of a supermarket. The information about the number of items sold and the sales revenue are on a weekly basis starting in September 1994. There are important influences on the sales that should be taken into consideration: advertising campaigns sometimes combined with temporary price reductions; holidays shorten the opening hours; the season has an effect on the sales of

the considered items.

We take the sales information, prices and advertising campaigns from the cash registers and the marketing team of the supermarket. The holidays are calculated. For the season information we use the time series of the turnover sum in DM of all items of this product group as an indicator. Its behavior over a term of 19 months is shown in figure 1.



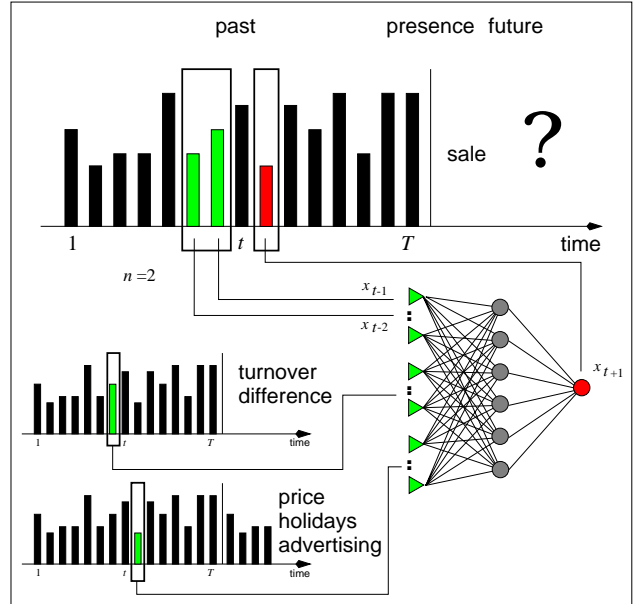
**Figure 1. Turnover sum in DM of the product group September 1994 to March 1996**

We use feed-forward multilayer perceptron (MLP) networks with one hidden layer together with the back-propagation training method. In order to predict the future sales the past information of  $n$  recent weeks is given in the input layer. The only result in the output layer is the sale for the next week.

Due to the purchasing system used in the supermarket there is a gap of one week between the newest sale value and the forecasted week. In addition the pricing information, advertising campaigns and holidays are already known for the future, when the forecast is calculated. This information is also given to the input layer as shown in figure 2.

### 3. Preprocessing the Input Data

An efficient preprocessing of the data is necessary to input it into the net. In general it is better to transform the raw time series data into indicators that represent the underlying information more explicitly. Due to the sigmoidal activation function of the back-propagation algorithm the sales information must be scaled to  $]0, 1[$ . The scaling is necessary to support the back-propagation learning algorithm [4]. We tested several scalings ( $z_t$ ) for the sale and the turnover time



**Figure 2. Input and output of the MLP**

series  $x = (x_t)$ :

$$z_t = \frac{x_t - \min(x)}{\max(x) - \min(x)} \cdot 0.8 + 0.1 \quad \text{resp.}$$

$$z_t = \frac{x_t - \mu}{c \cdot \sigma} + 0.5$$

where  $\min$  and  $\max$  are the minimum and maximum values of time series  $x$  and  $\mu$  and  $\sigma$  are the average and the standard deviation.  $c$  is a factor to control the interval of the values.

For the prices the most effecting indicator is the price change. So the prices are modeled as follows:

$$pr_{it} := \left\{ \begin{array}{l} 0.9 : \text{price increases} \\ 0.0 : \text{price keeps equal} \\ -0.9 : \text{price decreases} \end{array} \right\} \text{ within week } t$$

For both the time series of holidays and advertising campaigns we tested binary coding and linear aggregation to make them weekly. Their indicators are:

$$y_t := \left\{ \begin{array}{l} 0.9 : \text{if there is a holiday resp.} \\ \quad \text{advertising within week } t \\ 0.0 : \text{otherwise} \end{array} \right\} \text{ resp.}$$

$$y_t := \frac{\text{number of advertising resp. holidays within } t}{6}$$

(normalized number of special days within week  $t$ )

## 4. Experimental Results

To determine the appropriate configuration of the feed-forward MLP network several parameters have been varied:

- modeling of input time series
- width of the sliding time window
- the number of hidden neurons
- interval of the initial random weights
- training rate and momentum
- number and selection of validation patterns
- dealing with overfitting

To evaluate the efficiency of the neural network approach several tests have been performed. We compare the prediction error of a naive (“Naive”) and a statistical prediction method (“MovAvg”) to the successive prediction by neural networks (“Neural”).

### 4.1. Naive Prediction

The naive prediction method uses the last known value of the time series of sales as the forecast value for the future. In our terms:  $\hat{x}_{t+1} := x_{t-1}$ . This forecasting method is often used by the supermarket’s personnel.

### 4.2. Statistical Prediction

The statistical method is currently being used by the supermarket’s headquarters to forecast sales and to guide personnel responsible for purchasing. It calculates the moving average of a maximum of nine recent weeks, after these sale values have been filtered from exceptions and smoothed.

### 4.3. Neural Prediction

We reached good results for  $n = 2$  recent values of the sale time series in the sliding window. The other inputs are, one neuron each: both the difference of the sale ( $x'_t = x_t - x_{t-1}$ ) and the turnover of the whole group of items for the last week and the holiday, advertising and pricing information for the week to be predicted.

Thus, for each item a net with 7 input neurons and 4 hidden neurons is trained for a one week ahead forecast with a gap of one week. We reached better results with the binary scaling for the holiday and advertising time

series and the  $\sigma$ - $\mu$ -scaling for sales. The learning rate of the back-propagation algorithm was set to 0.3 with a momentum of 0.1. The initial weights were chosen from  $[-0.5, 0.5]$  by chance. The training was validated by 12 patterns and stopped at the minimum error.

### 4.4. Comparison of Prediction Techniques

The results of the forecasting accuracy are calculated for the successive prediction of the 22 weeks 44/1995 to 13/1996. In this weeks there are influences of many campaigns and Christmas holidays.

To measure the error the root mean squared error (RMSE) is divided by the mean value (Mean) of the time series. This is done for all the 20 items and the average value is show in table 1. Theil’s  $U$ -statistic is calculated as well.

**Table 1. Measuring forecasting accuracy, average of 20 items**

Accuracy Measure	Neural	MovAvg	Naive
RMSE/Mean	0.84	1.01	1.16
Theil’s $U$	0.76	0.92	1.00

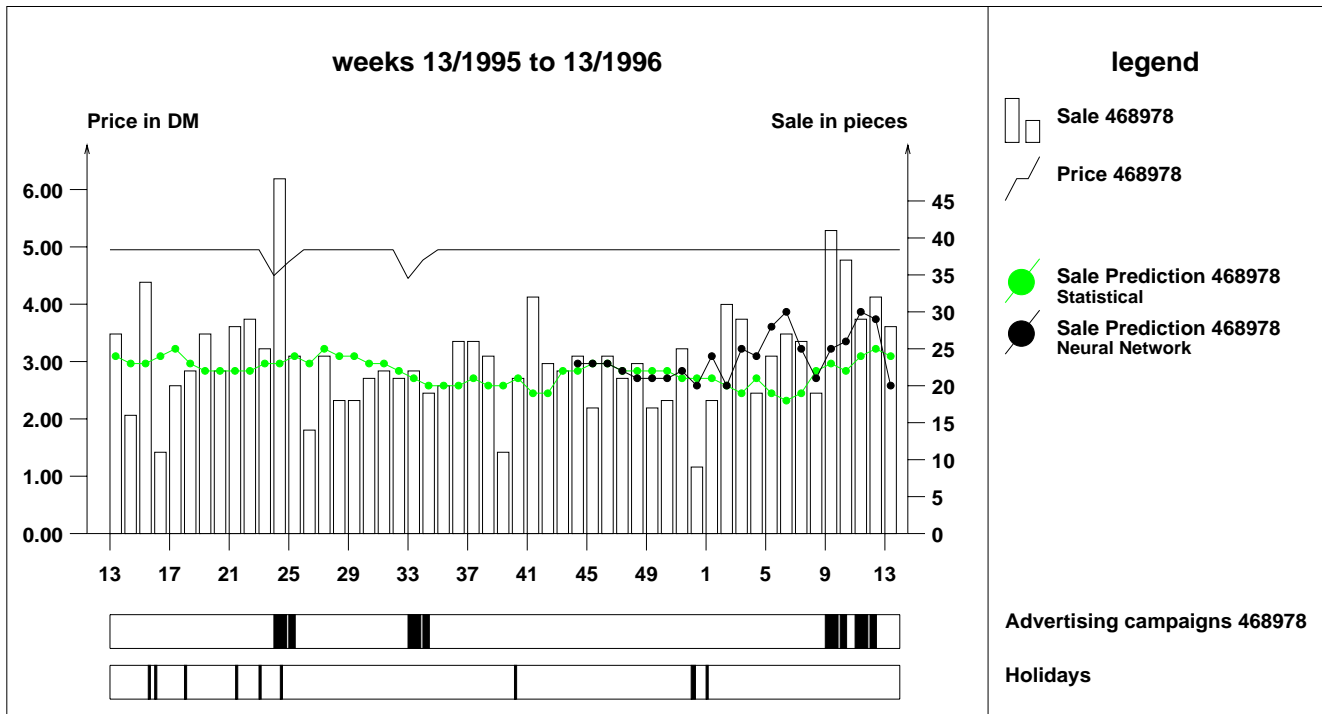
Based on the information in table 1 the naive approach is outperformed by the two other methods. For 18 of the 20 items the prediction by the neural network is better than the statistical prediction method.

A close inspection of the times series favored by the statistical approach shows that these are very noisy without any implicit rules that could be learned by the neural network. Especially one of these items has an average weekly sale of less than 4 items.

Figure 3 shows the predicted values for item 468978 calculated by the statistical and neural approach. The price, advertising and holiday information is included in this figure.

## 5. Conclusions and Future Research

For a special group of items in a German supermarket neural nets have been trained to forecast future demands on the basis of the past data augmented with further influences like price changing, advertising campaigns and holiday season information. The experimental results show that neural nets outperform the naive and statistical approaches that are currently being used in the supermarket.



**Figure 3. Comparison of sale prediction for an item by statistical and neural approach**

Our procedure preprocesses the data of all kinds of time series in the same manner and uses the same network architecture for the prediction of all 20 time series of sales. The parameter optimization is based on all of the time series instead of on one special item.

The program runs as a prototype and handles only a small subset of the supermarket's inventory. Future work will concentrate on the integration of our forecasting tool into the whole enterprise data flow process. Since a huge number of varying products have to be managed a selection process has to be installed that discriminates between steady time series suitable for conventional methods and chaotic candidates which will be processed by neural nets.

The prototype is part of a forecasting system that is able to take the raw data, do the necessary preprocessing, train the nets and produce an appropriate forecast. The next steps will be the development of additional adaptive transformation techniques and methods to test the significance of inputs which can be used to reduce the complexity of the nets.

## References

- [1] K. Chakraborty, K. Mehrotra, C. Mohan, and S. Ranka. Forecasting the behaviour of multivariate time series using neural networks. *Neural Networks*, 5:961–970, 1992.
- [2] B. Freisleben and K. Ripper. Economic forecasting using neural networks. In *Proceedings of the 1995 IEEE International Conference on Neural Networks*, volume 2, pages 833–838, Perth, W.A., 1995. IEEE.
- [3] A. Refenes, M. Azema-Barac, L. Chen, and S. Karousos. Currency exchange rate prediction and neural network design strategies. *Neural Computing & Applications*, 1(1):46–58, 1993.
- [4] H. Rehkugler and H. G. Zimmermann. *Neuronale Netze in der Ökonomie (in German; Neural Networks in Economics)*. Verlag Vahlen, München, 1994.
- [5] R. Rojas. *Neural Nets*. Springer, 1996.
- [6] E. Schöneburg. Stock price prediction using neural networks: An empirical test. *Neurocomputing*, 2(1), 1991.
- [7] V. R. Vemuri and R. D. Rogers. *Artificial Neural Networks – Forecasting Time Series*. IEEE Computer Society Press, 1994.
- [8] A. S. Weigend and N. A. Gershenfeld. *Time Series Prediction: Forecasting the Future and Understanding the Past*. Addison-Wesley, 1994.