# Neural Networks for Financial Market Prediction

C.H. Chen
Electrical and Computer Engineering Dept.
University of Massachusetts Dartmouth
North Dartmouth, MA 02747-2300 USA

## Abstract

The use of neural networks in financial market prediction presents a major challenge to the design of effective neural network predictors and classifiers. In this paper, we have examined several neural networks to evaluate their capability in prediction and in trend estimation which is treated as a classification problem. The networks considered are the backpropagation trained network (BPN), the general regression neural network (GRNN), the class-sensitive neural network (CSNN), and the conjugate gradient trained network (CGNN). It is concluded that CSNN is among the best performing networks in both prediction and trend estimation. All major indicators are evaluated by the neural networks. It is found that the use of good indicators like the rate of change, momentum, moving average, etc. can lead to about 5% improvement over the case that no indicator is used. Momentum computed from the previous 14 day data is the best single indicator. The complexity of the financial market probably explains why the large number of indicators cannot provide any significant improvement in network classification.

## 1. Introduction

The progress in time series analysis and statistical pattern recognition has generated a lot of interest in the last 30 years to predict the financial market time series data. The success has been very limited because neither approach has been able to model the stock market data adequately. The rapid progress of artificial neural network (hereafter called neural network) study in the last few years has generated a lot of interest in the short-term stock market prediction by using neural networks (see e.g. [1-6]). Neural networks derive the prediction and classification capabilities from the training data that is part of the same data set. The results thus are not dependent on sometimes unrealistic model assumption of the data. Of course different neural networks do not perform the same. After a comparison of several feedforward networks, we have concluded that the class-sensitive neural network (CSNN) [7] is one of the best performing networks in both prediction and classification. Both performance comparison and the CSNN learning process as well as architecture will be examined in this paper.

A recommendation of the study is that the use of a dedicated neural network system for short-term stock market prediction is feasible. Such system will help the human operator in making intelligent decisions and predictions, but, like many artificial intelligent systems developed so far, it cannot replace the experienced human operator.

## 2. Data Characteristics

The initial data as provided by Morningside N.A. consists of four sets, each corresponding to one full year of daily data. Data sets A and B are for the stock market of 1986 and 1991 respectively. Data sets C and D are for the currency market of 1990 and 1992 respectively. Each data set has for each day the open, high, low and closing prices. A plot of the data set A is shown in Fig. 1 which contains four time series. The second set of data is the Dow Jones Averages for 1980 including high, low and closing prices for each day. A statistical analysis of the initial data sets shows that a proper auto-regressive model representation of the data requires only an order of 2 to 15. The long term periodicity of the data is not evident.

## 3. Comparison of Prediction Performances

The prediction performance is measured by the correlation between the predicted value and the actual value, rxy, and by the prediction error between the predicted and actual values. The prediction error, represented as root mean square error, rmse, is not normalized. The correlation is normalized so the largest correlation is 1. For a

0-7803-1901-X/94 $4.00 ©1994 IEEE

good prediction, correlation should be large and the rmse should be small.

The tabulation in Table 1 shows a prediction performance comparison of linear prediction (LP) based on the autoregressive model of order 15, backpropagation trained network (BPN), general regression neural network (GRNN) [8], and the class-sensitive neural networks (CSNN). The linear prediction is worse than the neural networks in both correlation and the rmse. BPN has a larger correlation than the GRNN but a smaller rmse than GRNN, but the difference between the two is small. Though CSNN is designed to minimize the classification error, it has the best prediction performance (less than 3%). The results are fairly consistent with other data sets. In other words the above conclusion is valid for all data sets studied. A comparison has also been made between the use of two day's data and the use of three day's data in prediction, for prediction of 1,2, and 3 days in advance, based on CSNN. A general conclusion is that the use of two day's data is slightly superior.

## 4. Comparison of Trend Estimation (Classification) Performance

By trend estimation we mean the decision on whether the next day's price is going up (or unchanged) or down. We have done trend estimation for data sets A,B,C, and D using both CSNN and the conjugate gradient trained neural network ((CGNN)[9] with 12 features including two day's raw data and the MACD (moving average convergence and divergence) indicator. The reason CGNN is chosen is because it has a comparable classification performance as CSNN from our experience with the sonar data. Other networks such as BPN and GRNN may have less classification performance than CSNN or CGNN. For the stock market data, CSNN is shown to be better than CGNN in almost all cases considered.

Another experiment is on the Dow-Jone Average Data (the second data set mentioned earlier) trend estimation. Again it is found that the closing price prediction is not as good as the high and low price predictions. A further experiment is performed by including volume data in trend estimation for the Dow-Jones Average Data using CSNN. By using the volume data, the closing trend prediction is improved. The performance on high and low trend prediction degrades for unknown reason. One explanation is that the volume information may be redundant for high and low estimation. It is however reasonable to conclude that in predicting a specific stock market, it is important to consider more economic indices, such as the Dow Jones Industrials, stock volumes, interest rate and some other indices, rather than just the stock value and the related indicators.

## 5. Effect of Indicators on Network Trend Estimation

Our perception that market indicators can help with the stock market prediction is only partially correct from the neural network study. There can be a lot of overlapped information among the indicators. In fact the pattern recognition theory states that the best features are those which have independent information. Thus it is not unexpected that the best available improvement is only about 5% in trend estimation with the use of best combination of indicators.

Table 2 lists the performances of the best indicator combinations for the closing price trend estimation with CSNN and the backpropagation trained network. For the data set D, the addition of economic indices from the second set of data mentioned earlier is considered. By including such information, it is found that the performance may improve or degrade depending on the type of indices used. Only for dollar index (DX) we can observe a consistent improvement for all three prices. A combination of dollar index with Eurodollar (ED) or with Standard & Poor (SP) also shows a consistent improvement in all three prices.

## 6. CSNN Learning Process and Architecture

To design a network and to test the network are really two different processes. Normally a given data is divided into a training set and a test set.

The network is trained until the mean square error reaches a steady but small value which hopefully is the absolute minimum, though the network can be trapped to a local minimum. There are ways to avoid the local minimum but this subject is not discussed here. A network that is trained to an absolute minimum does not guarantee the best classification or generalization capability. In CSNN we constantly examine the network's classification performance. For an objective evaluation of network performance, we have devised the following strategy. First we construct a training set of 100 points. Then we construct two testing sets. The first testing set (30 points) is used to evaluate the network, the second (10 or 30 points) is used for real test. Train the network until the first testing set gets to the highest correct recognition. Here we only consider the training passes from 200 to 1000. When the training passes exceeds 1000, the testing performance may not be improved or may even degrade while the training performance can still improve.

As for the CSNN architecture, one hidden layer for each subnet usually works quite well and the number of input neurons is the same as the number of input features. A subnet is used for each class while all classes share the same input. The number of hidden layer neurons for each class is chosen between 20 and 32 and is greater than the number of input neurons. There is one output neuron for each class and thus the total number of output neurons is the same as the number of classes considered. The learning constant chosen is 0.05. The period of adjusting learning constant is 20. It takes about 200-400 passes to finish the training procedure, with a training time of 6-20 minutes in an IBMPC486 with 33 MHz. A software package developed under Window 3.1 based on Microsoft C++ has been completed by us.

## 7. Concluding Remarks

After many years of efforts in using time series analysis, statistical pattern recognition and artificial intelligence, neural networks now appear now to be the technology that offers effective computer-based financial market prediction. Even though the best neural network results still have a lot to be desired the key is in better neural networks that are realizable in hardware. The possibility of combining neural network with knowledge based expert systems is another potential area of research. Also a good graphical display environment is important for human interpretations of compute results.

## References


1. H. White, "Economic prediction using neural networks: The case of IBM daily stock returns", Proc. of the IEEE International Conference on Neural Networks (ICNN'88), San Diego, CA, vol. II, pp. 451-459, 1988.
2. C. Chang, B. Sheu and J. Thomas, "Multi-layered back-propagation neural networks for finance analysis", Proc. of World Congress on Neural Networks (WCNN), vol. I, pp. 445-450, Portland, OR, July 1993.
3. C. Epstein, "The thinking man's machine", Futures and Options World, pp. 51-53. Sept. 1992.
4. T. Kimoto, et.al., "Stock market prediction system with modular neural networks", Proc. of ICNN 1990.
5. K. Kamijo and T. Tanigawa, "Stock price pattern recognition - a recurrent neural network approach", Proc. of ICNN 1990.
6. X. Zhang and J. Hutchinson, "Simple architectures on fast machines: practical issues in nonlinear time series prediction", in "Time Series Prediction: Forecasting the Future and Understanding the Past", eds. A.S. Weigend and N.A. Gershenfeld, Addison-Wesley, 1993.
7. C. H. Chen and G.H. You, "Class-sensitive neural network", Neural, Parallel and Scientific Computing, vol. 1, no. 1, March 1993.
8. D.F. Specht, "A general regression neural network", IEEE Trans. on Neural Networks, vol. 2, no. 6, pp. 568-576, Nov. 1991.
9. C. H. Chen and H. Lai, "An empirical comparison of gradient descent and conjugate gradient backpropagation trained neural networks", OCEANS'92, Newport, RI, Oct. 1992.

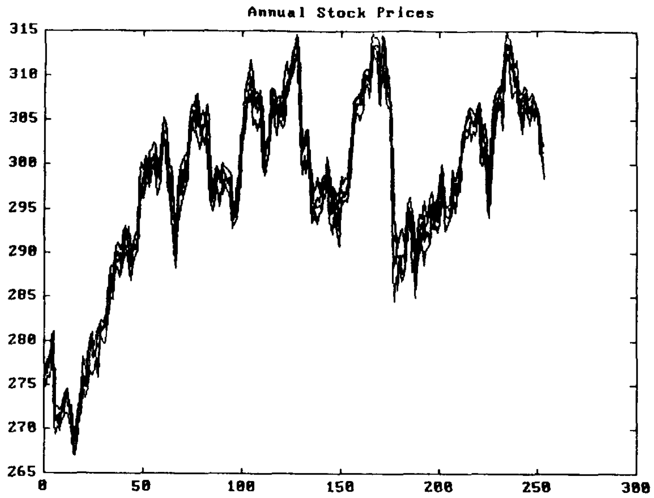


Figure 1 Plot of the four time series of data set A.

The following tabulation shows a comparison of linear prediction (LP), BP, GRNN, and CSNN in the one-day prediction of the "A" time series of the stock market data.

| | Open | High | Low | Close |
|---|---|---|---|---|
| LP rxy = | 0.896 | 0.925 | 0.904 | 0.891 |
| rmse = | 3.274 | 2.674 | 3.140 | 3.295 |
| BP rxy = | 0.976 | 0.971 | 0.955 | 0.946 |
| rmse = | 2.177 | 2.360 | 2.980 | 3.266 |
| GRNN rxy= | 0.975 | 0.953 | 0.923 | 0.885 |
| rmse = | 1.597 | 2.066 | 2.793 | 3.271 |
| CSNN rxy= | 0.992 | 0.973 | 0.904 | 0.891 |
| rmse = | 1.012 | 1.630 | 2.342 | 2.926 |

Table 1 Comparison of prediction performances.

**CSNN performance:**

| indicator | a.csv(closing) | c.csv(closing) |
|---|---|---|
| no | 59.20% | 56.45% |
| RSI, MM | 63.20 | 60.00 |
| ROC, RSI, MM | 62.40 | 65.60 |
| MA, ROC, RSI, MM | 63.20 | 64.80 |

**BPN performance:**

| indicator | a.csv(closing) | c.csv(closing) |
|---|---|---|
| no | 50.86% | 52.72% |
| RSI, MM | 54.76 | 51.88 |
| ROC, RSI, MM | 54.74 | 53.56 |
| MA, ROC, RSI, MM | 55.82 | 54.68 |

Notes: MA=Moving Average; ROC=Rate of Change;
RSI=Relative Strength Index; MM=Momentum

Table 2 Performance of best indicator combinations.