

# One Day Prediction of NIKKEI Index Considering Information from Other Stock Markets

Marcin Jaruszewicz and Jacek Mańdziuk

Faculty of Mathematics and Information Science, Warsaw University of Technology,  
Plac Politechniki 1, 00-661 Warsaw, POLAND  
jaruszewicz@data.pl, mandziuk@mini.pw.edu.pl

**Abstract.** A task of a stock index prediction is presented in this paper. Several issues are considered. The data is gathered at the concerned stock market (NIKKEI) and two other markets (NASDAQ and DAX). The data contains not only original numerical values from the markets but also indicators pre-processed in terms of technical analysis, i.e. the oscillators are calculated and the structures of a value chart are extracted. Selected data is input to a neural network that is functionally divided into separate modules. The prediction goal was next day opening value of Japanese stock market index NIKKEI with consideration of German and USA stock markets' indexes. The average prediction error on the test set equals 43 points and the average percentage prediction error is equal to 0.27% while the average index volatility equals 0.96%.

## 1 Introduction

Prediction of a stock market index is a very complex task due to the lack of the autocorrelation of an index value changes even in a period of one day [1]. Despite there is only one variable to be traced, many derived variables are considered. Also not measurable incidents like economic or political situation, catastrophe or war may cause the change of a stock market index [1]. Except for the above variables there is also some information provided by technical analysis [2], which should be considered. Prediction system presented in the paper uses all the above different types of information. The most important stage of a prediction process is a complex data analysis and pre-processing. In the next parts the data pre-processing and neural network are described in detail. In our experiment the first task was prediction based solely on historical data from one stock market only. In the next step the data from two other stock markets was added [3] and the most important input variables were selected [4]. Finally, based on the above two steps the network with separated regions (by incomplete weights) was created [5, 6].

## 2 Data Pre-processing

The data was collected on a daily basis. For each day the sample was composed of the date, the opening, closing, highest and lowest values.

Having values of a stock market index statistical variables were calculated. First of all, percentage changes in the period of 1, 5, 10 and 20 days in the past, resp. were evaluated. Weighted moving averages for the last 5, 10 and 20 days, resp. were also calculated:

$$mov(v, k) = \frac{\sum_{i=0}^{k-1} (k-i)v_{t-i}}{\sum_{i=0}^{k-1} (k-i)} \quad (1)$$

where  $v_t$  is an index value in time  $t$ ,  $k$  is the length of a period. Finally technical analysis data was evaluated. Three different oscillators were calculated [7]. Proper definitions are described below.

$$MACD = mov(C, 10) - mov(C, 20), \quad MACD_{SIGNAL} = mov(MACD, 5) \quad (2)$$

where  $C$  is the index closing value. Signals are generated in the intersections of  $MACD$  and  $MACD_{SIGNAL}$ .

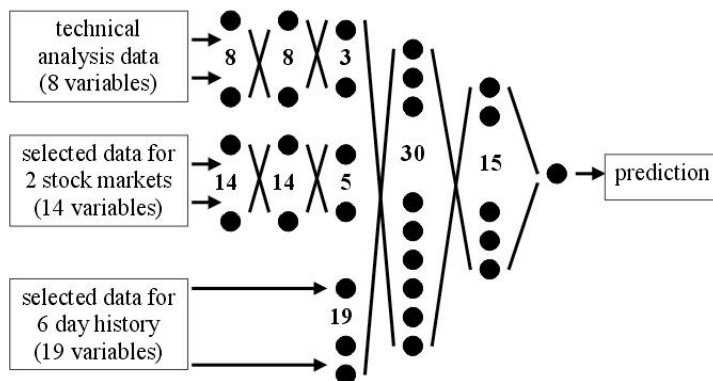
$$WILLIAMS(n) = \frac{\max(H, n) - C}{\max(H, n) - \min(L, n)}(-100) \quad (3)$$

where  $n$  is a period of time (set to 10 days in the paper),  $\max(H, n)$  is the highest index value in this period,  $C$  is index closing value. The final oscillator called ‘‘Two averages’’:

$$2AVG = mov(C, 5) - mov(C, 20) \quad (4)$$

generates signals according to its sign changes.

Signals generated by the above oscillators were added to the learning data. Also structures known in technical analysis [7] were extracted. An extraction is based on an analysis of the shape of the chart of the index value. These structures forecast change or continuation of a trend. Information of structures and signals of trend changes generated according to them were also included in the learning data. All ‘‘buy’’ signals were coded as 0.8 value and all ‘‘sell’’ signals as  $-0.8$ . Information about signal was copied into samples through the next 5 days with values linearly decreasing in time. It was done to prevent the existence of too many samples with no signal value. Anyway, in practice, the signal is usually correct during the next few days after its appearance. Signals cumulate when appear more than once for a single record. All above data was generated for three aforementioned stock markets. After all the above calculations 4 399 records were created covering dates 1985/10/30 – 2003/06/02. 400 records out of the above number were randomly selected as the testing set. This data was not used in the training phase. Since the stopping condition depended only on the number of training epochs (there were 4,000 epochs for single, preliminary networks and 10,000 ones for final one), no validation set was used in the experiment.



**Fig. 1.** Structure of the final network with separated regions for different input data types. Values in parentheses denote the sizes of subsequent layers - from the input layer (the left value) to the output one (the right value).

### 3 System Architecture

In the past, different predictors were considered with a prediction of stock markets [8, 9]. In this paper we focus on data structures, so the prediction system was based on a simple multilayer perceptron network. In the first half of training iterations a backpropagation method with momentum was used, and for the rest of iterations the standard backpropagation method was applied. The input data was normalized and scaled to the range  $(-1, 1)$ . In all neurons sigmoidal activation function was used.

Initially, all weights in the network were randomly chosen from the range  $(-0.01, 0.01)$ . Due to a large number of potential input variables, before the final prediction some tests of variables importance were made. First of all prediction using only one considered stock market and its past 6 days history was done by neural network of size  $108 \times 50 \times 1$  with one hidden layer. In the input layer 108 values representing the data from the target stock market were presented. In the output layer the next day's opening value was predicted.

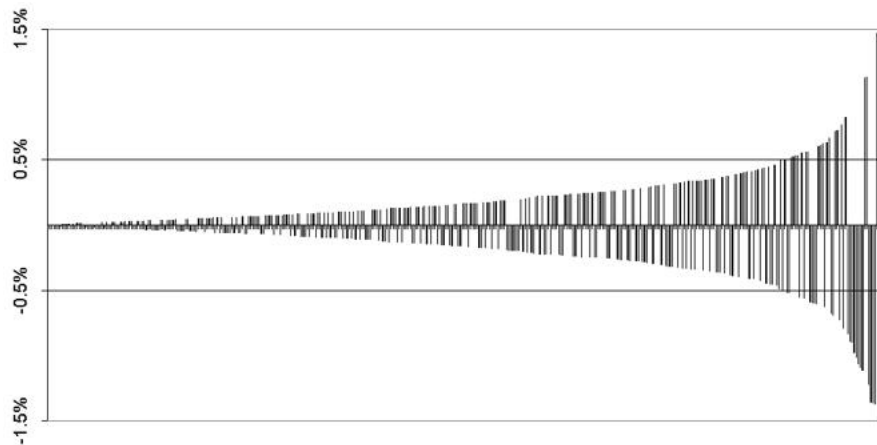
In the next test the data from three stock markets was considered. Together there were 55 variables describing situation on these markets during one day of prediction. The size of a hidden layer was equal to 30 and the output layer was composed of one neuron. Both tests aimed at selection of the most relevant inputs for the final prediction system.

The final prediction was made by a network with separate regions for each input data type. The first region processed the most relevant data (selected in preliminary stage) from two stock markets other than the market concerned in prediction. The second region processed the most important technical analysis data from target stock market. The most relevant data from 6 day history was given to the network without any further preprocessing. A structure of the final neural network is presented in Fig. 1.

## 4 Results

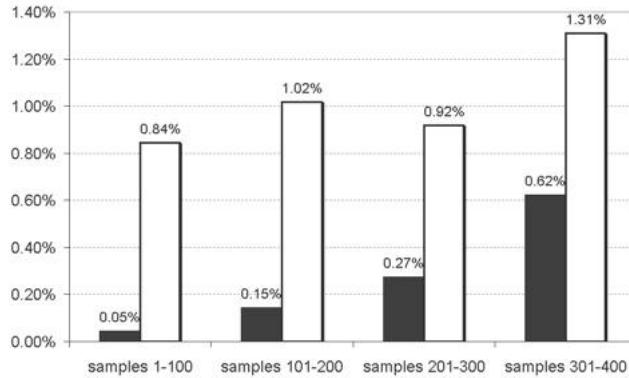
Prediction based on 6 day history data from one market only appeared to be a difficult task for the network. After 4 000 iterations the average error on test samples was equal to 4.85%, which is much more than the average volatility of the NIKKEI index, which is 0.96%. During this test the importance of particular input variables was measured. Using the chart of sum of absolute values of outgoing weights for each neuron of the input layer, the number of 19 out of 108 variables were selected for the final prediction. Much better results were obtained after adding the information from two other stock markets except historical data. The average percentage error was equal to 1.96%. The input data was from a single day for three stock markets (target and two additional) in the number of 55. After the learning process 14 variables from two non-target markets were selected for the final experiment.

In the final experiment, except for the historical data and the data from two other markets also a technical analysis data was used. Therefore the final selection of input variables was the following (cf. Fig. 1). In the **historical data** 19 inputs, i.e. opening, closing, lowest values and moving average for 20 days for **day t**; closing, lowest values and moving averages for 20 days for **days t-1, t-2, t-3** and **t-4**; lowest value as well as moving averages for 20 and 5 days for **day t-5**. Among **other markets' data** 14 inputs, i.e. the opening, closing, highest and lowest values as well as moving averages for 5, 10 and 20 days for two non-target stock markets. And in **technical analysis data**, (8 inputs): closing value, structure type, MACD line, MACD signal line, Williams as well as 3 signals generated by oscillators: MACD, Williams, Two Averages.



**Fig. 2.** A structure of a percentage prediction error (Y-axis) on test samples (X-axis) in the final experiment. Samples are ordered from the least to the highest module of a value of a percentage error.

After 10 000 iterations the average percentage error was equal to 0.27%. Prediction error accomplished in the final experiment is more than three times lower than the average volatility of NIKKEI index, i.e. 0.96% (with maximum equal to 13.23%).



**Fig. 3.** A cumulated prediction error in the final experiment - average percentage values (Y-axis) of an error (black boxes) and index volatility (white boxes) for the sets of samples (X-axis). The size of each set is equal to 100.

The structure of the error is presented in Fig. 2. Samples are sorted by increasing value of module of a value of a percentage error. The manifest observation is that the value of an error doesn't grow linearly.

In Fig. 3. the average values of an error are shown for different set of test samples together with the average real volatility of an index value. Each set consist of 100 samples. The first 100 test samples (with the lowest error) have the average error of 0.05%. Results for the next sets of samples are 0.15% and 0.27%, respectively. An error value for the last 100 samples (with the highest error) equals 0.62%. The volatility of an index value changes from 0.84% to 1.31%. Despite there is no strong correlation between an error value and volatility of an index, the error value grows together with the volatility value. Samples with a rapid index change during a day are a little more difficult for prediction than samples with the low volatility.

Similarly to preliminary experiments the relative importance of particular inputs was measured based on the sum of modules of all outgoing weights divided by the number of these weights. It appeared that the most important was the data from two other (non-target) stock markets with the average value of 0.33. Values for historical and technical analysis data were equal to 0.22 and 0.01, respectively.

After removing the entire region devoted to technical analysis data (see Fig. 1) the average percentage error increased to 0.37%.

## 5 Conclusions

A way from a simple neural network to a network based on separate regions producing better forecast was presented. Promising test results were obtained. The percentage error equal to 0.27% is a good starting point for further investigation, especially in comparison with an average volatility of the index value being at the level of 0.96%.

Further research will focus on usefulness of technical analysis data. The main point is the observation that oscillators have very unstable impact on proper buy/sell signal generation. Each oscillator is useful only in a short period of time. In that case weights between technical analysis region and the main network should change respectively. Probably the same observation is correct for other variables.

Another field of experiments is research on a structure of the network. Simple model-based input variables selection described in this paper will be extended by model-free methods [10, 11].

A far-reaching goal is a buy/sell signal prediction, which is more complex task than a value prediction due to the necessity to predict a trend in a horizon longer than one day.

## References

1. Mantegna, R., Stanley, E.: An introduction to econophysics. Correlations and complexity in finance. Cambridge University Press (2000)
2. Dempster, T.P., et al.: Computational learning techniques for intraday fx trading using popular technical indicators. *IEEE Transactions on Neural Networks* **12** (2001) 744–754
3. Podding, T., Rehkegler, H.: A “world” model of integrated financial markets using artificial neural networks. *Neurocomputing* **10** (1996) 251–273
4. Chenoweth, T., Obradović, Z.: A multi-component nonlinear prediction system for the s&p 500 index. *Neurocomputing* **10** (1996) 275–290
5. Fu, H.C., Lee, Y.P., et al.: Divide-and-conquer learning and modular perceptron networks. *IEEE Transactions on Neural Networks* **12** (2001) 250–263
6. Khotanzad, A., Elragal, H., et al.: Combination of artificial neural-network forecasters for prediction of natural gas consumption. *IEEE Transactions on Neural Networks* **11** (2000) 464–473
7. Murphy, J.: *Technical analysis of the financial markets*. New York Institute of Finance (1999)
8. Tony Gestel, J.S., et al.: Financial time series prediction using least squares support vector machines within the evidence framework. *IEEE Transactions on Neural Networks* **12** (2001) 809–820
9. Peter Tino, C.S., et al.: Financial volatility trading using recurrent neural networks. *IEEE Transactions on Neural Networks* **12** (2001) 865–874
10. Back, A., Trappenberg, T.: Selecting inputs for modeling using normalized higher order statistics and independent component analysis. *IEEE Transactions on Neural Networks* **12** (2001) 612–617
11. Refenes, A., Holt, W.: Forecasting volatility with neural regression: A contribution to model adequacy. *IEEE Transactions on Neural Networks* **12** (2001) 850–864