Forecasting the dynamics of financial time series based on neural networks

V A Ivanyuk¹², N M Abdikeev¹ and A D Tsvirkun³

¹ Department of Data Analysis and Machine Learning, Financial University under the Government of the Russian Federation, 49 Leningradsky Prospekt, 125993, Moscow, Russia
² Department of Higher Mathematics, Bauman Moscow State Technical University, Moscow, Russia
³ V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Russia

E-mail: VAIvanyuk@fa.ru

Abstract. Forecasting is one of the high-demand data mining problems, but also a very difficult one. The difficulties of forecasting are associated with insufficient quality and quantity of input data, the changes in the environment where the process takes place, and the impact of subjective factors. A forecast always implies some margin of error, which depends on the forecast model used and the completeness of the input data. Methods based on neural networks are the most relevant and highly-demanded techniques today. Neural networks are great for finding accurate solutions in an environment characterized by complex or fragmented information. In the field of finance and economics, the values of time series parameters can be more accurately modelled using neural analysis methods. Artificial neural networks have more common and flexible functional forms than statistical methods. They can generalize information and provide a qualitative forecast under conditions of uncertainty and crisis. The article proposes a forecasting model based on a neural network that can predict the price of a financial asset in a well-defined time interval. Ten technical indicators are used as input signals, and the closing price of the next period is used as an output signal.

1. Introduction

Within the framework of the analysis of perspective models and forecasting methods, several works can be highlighted. For example, Hi’ovská K and Koncz P [1] compare a nonlinear ARIMA-model and an artificial neural network, which leads to the conclusion that AI-based methods show a more accurate result.

The paper by Yajiao Tang and Junkau Ji [2] shows that models based on artificial intelligence are more accurate on real data.

The authors Pyo S, Lee J, Cha M and Jang H [3] reinterpreted the financial forecasting paradigm, using indicators as sources of basic market information and a continuous input classification model as the neural network architecture.

Neural network-based models are the most popular structured methods for predicting time series [4, 5]. The popularity of this tool is due to their ability to set non-linear dependencies and to adapt [6, 7].

Artificial neural networks (ANNs) are computing systems that partially resemble biological neural networks. Such systems «learn» to perform tasks by looking at examples.
ANN is a model based on a collection of connected blocks or nodes called «artificial neurons». Each connection, like synapses, can transmit information, a «signal», from one artificial neuron to another. An artificial neuron that receives a signal can process it and then signal additional artificial neurons connected to it.

In common ANN implementations, the signal at the connection between artificial neurons is a real number, and the output of each artificial neuron is computed by some nonlinear function of the sum of its inputs. Artificial neurons usually have a weight that adjusts during the learning process [8, 9]. The weight increases or decreases the signal strength at a connection. Artificial neurons can have a threshold such that a signal is sent only if the aggregate signal crosses this threshold. Typically, artificial neurons are assembled into layers. Different layers can perform different types of transformations on their inputs. Signals move from the first layer (input) to the last layer (output), possibly after passing through the layers multiple times.

The original goal of the ANN approach was to solve problems in the same way that the human brain does. However, over time, the focus shifted to specific tasks. Artificial neural networks are used for various tasks, including computer vision, speech recognition, and machine translation [10, 11, 12]. Neural networks are actively gaining popularity in forecasting financial time series, as they show good results on time series with high volatility [13, 14].

2. Methods of neural network training

The main concern in neural network modelling is the selection of a neural network training algorithm, which determines the speed and quality of construction [15, 16]. The following basic methods of neural network training can be distinguished.

Gradient descent method

The method assumes finding the local minimum of the function by moving along the gradient. The idea of the method is to optimize the direction of movement, which is set by a gradient (1)

\[ x[k+1] = x[k] - \lambda[k] \nabla f(x[k]) \]  

This method, in its turn, is classified as follows depending on the stepsize:

- gradient descent with constant stepsize;
- gradient descent with step splitting;
- steepest descent (2).

\[ \lambda[k] = \arg\min_{\lambda} f(x[k] - \lambda[k] \nabla f(x[k])) \]  

One of the following conditions can be used as a criterion for stopping the function optimization process (3-4):

\[ ||x[k+1] - x[k]|| \leq \varepsilon \]  

\[ ||f(x[k+1]) - f(x[k])|| \leq \varepsilon \]

The Levenberg – Marquardt algorithm.

This optimization method aims at the least-squares minimization (5). By its nature, it is similar to the gradient descent method.

\[ (J^T J + \mu I)d = -J^T r, \mu \geq 0 \]  

The conjugate-gradient method.

This is an iterative method of optimization in multidimensional space that solves the problem of optimizing a (quadratic) function in a finite number of steps (6):

\[ a_k = \arg\min_{a_k} F(x_{k-1} + a_k p_k) \]  

The next stage is to choose the neural network architecture, for example:

- nonlinear autoregressive with external input;

To generate a network based on this architecture, one must consider the following arguments:

- input delay vector;
- feedback delay vector;
• the number of hidden neurons;
• training method.

- non-linear input-output (without feedback connections);

To generate a network based on this architecture, one must consider the following arguments:

• input delay vector;
• the number of hidden neurons;
• training method.

The next stage is to determine the number of hidden neurons. A couple of basic methods are listed below (where \( k \) is the number of hidden neurons; \( n \) is the number of neurons at the input; \( m \) is the number of neurons at the output):

1) \( k = \sqrt{nm} \)
2) \( k = m^{\frac{n}{m}} \)

It is a common practice to use iterative methods for selecting the number of neurons. First, a model with \( k \) hidden neurons is built, and subsequently this number changes (7).

\[
k = \arg\max_k R(x_k), \forall k: k \in N
\]  

(7)

Any artificial neural network has input and output signals. As an output signal, the model uses the closing values in the period \( t \), while the following financial asset performance indicators are used as input signals:

• opening price in the period \( t - 1 \);
• closing price in the period \( t - 1 \);
• maximum price in the period \( t - 1 \);
• minimum price in the period \( t - 1 \).

Thus, having input and output signals, it is possible to derive a functional dependence in the neural network (8)(11).

\[
\text{Close}_t = f(O_{t-1}; H_{t-1}; L_{t-1}; C_{t-1})
\]  

(8)

The function depends on the opening price, the highest price, the lowest price, and the closing price, respectively.

3. **Time series prediction using an exogenous artificial neural network**

Global trends show a change in the area of financial time series forecasting. Whereas previously retrospective data was used for forecasting, now large-scale research is being conducted in the field of forecasting, relying on the classification based on exogenous variables. The latter can be represented by technical indicators. Consider the following indicators.

3.1. **Moving average with a period of 10 days (10-day MA)**

The moving average can be seen as a function that, at a point, equals to the average value of the previous time interval of a certain length. The method is primarily used for smoothing the time series (9).

\[
MA(10) = \frac{1}{10} \sum_{i=0}^{n-1} p_{t-i}
\]  

(9)

3.2. **Weighted moving average with a period of 10 days (10-day weighted MA)**

This indicator is similar to the simple moving average as it is used for the same purposes, but the difference is that the method allows placing more weight on the more current time periods (10).

\[
VMA(10) = \frac{\sum_{i=0}^{n-1} V_{t-i} p_{t-i}}{\sum_{i=0}^{n-1} V_{t-i}}
\]  

(10)

The moving average has the form (11):
\[(f \ast g)(x) = \int f(y)g(x - y)dy = \int f(x - y)g(y)dy\] (11)

3.3. Momentum.
The main task of this indicator is to measure the amount of price change. The indicator is not averaging, so many consider it to be a leading signal. Momentum belongs to the basic tools (12). The number of periods is usually taken to be equal to five.

\[\text{momentum} = C_t - C_{t-n}\] (12)

3.4. Fast stochastic oscillator K%
Oscillators are a class of indicators that show the position of the current price relative to the price range for a given retrospective time interval (13). The main idea is that in an uptrend of the price, the closing price tends to close near the maximum. We can say that the indicator shows the discrepancy between the closing prices of the period under review and the retrospective periods.

\[\%K = \frac{C_t - L_n}{H_n - L_n},\text{where}\] (13)

\[C_t\] – the closing price;
\[L_n\] – the lowest price;
\[H_n\] – the highest price.

3.5. Slow stochastic oscillator D%
The indicator is a moving average to the stochastic oscillator K% with a short averaging period. In practice, different averaging mechanisms can be used, including exponential or weighted averages.

3.6. Relative strength index
The index shows the strength of the trend and the probability of its change. In technical analysis, it is one of the most common indicators due to its simplicity of interpretation. It is calculated using indicators of up (U) and down (D) price changes (14 - 15).

\[U = C_t - C_{t-1}; D = 0\] (14)

\[D = C_t - C_{t-1}; C = 0\] (15)

It should be noted that in a situation where the closing prices of two adjacent periods are equal, (U) and (D) are set to zero. The price change indicators are smoothed using a modified exponential average with the given period (N). In this way, the relative strength is calculated (16):

\[RS = \frac{EMA(N)F(x)[U > 0]}{EMA(N)F(x)[D > 0]}\] (16)

Based on RS, one can calculate the RSI as follows (17):

\[RSI = 100 - \frac{100}{(1 + RS)}\] (17)

The method is used to determine the divergence.

3.7. MACD Indicator
MACD stands for Moving Average Convergence/Divergence. First of all, the indicator is used to check and determine the strength and direction of the trend vector. Second, the indicator can be used to determine the reversal points. The calculation is based on the difference between linear and exponential moving averages (18).

\[MACD = EMA_S(P) - EMA_L(P),\text{where}\] (18)

\[EMA_S(P)\] – exponential moving average with a short price period.
\[EMA_L(P)\] – exponential moving average with a long price period.
Exponential smoothing is used to eliminate possible random fluctuations.
3.8. A/D indicator
The accumulation/distribution indicator takes into account the trading volume (19). A positive value of the index indicates a positive trend in the financial time series, while a negative value indicates a negative trend in the time series.

\[
\frac{A}{D} = \frac{C_t - O_t}{H_t - L_t} Q_t, \text{ where}
\]

\( C_t \) – the closing price of the current period;
\( O_t \) – the opening price of the current period;
\( H_t \) – the highest price of the current period;
\( L_t \) – the lowest price of the current period;
\( Q_t \) – the trading volume of the current period.

3.9. Percent range %R
The Williams’ Percent Range Indicator (%R) is a technical indicator the main purpose of which is to determine the overbought or oversold state of an asset, based on the current position of the closing price (20). The indicator is calculated based on the maximum and minimum over a certain period of time.

\[
R\% = \frac{\max_{i \in [t-n,t]} H_i - C_t}{\max_{i \in [t-n,t]} H_i - \min_{i \in [t-n,t]} L_i}, \text{ where}
\]

\( \max_{i \in [t-n,t]} H_i \) – the highest price over the period \([t - n; t]\);
\( \min_{i \in [t-n,t]} L_i \) – the lowest price over the same period.
\( C_t \) – the closing price of the current period.

3.10. Commodity channel index
The CCI indicator belongs to the category of oscillators (21). By its nature, the indicator is made to determine whether an asset is overbought or oversold. Unlike other indicators of the oscillator group, the commodity channel index characterizes the deviation of the current closing price from the moving average. The type of moving average and its period are set individually for each time series. The indicator is mostly in the range from -100 to +100, and when it goes beyond that range, an oversold/overbought signal occurs.

\[
CCI = \frac{M_t - SM_t}{0,015 \times D_t}, \text{ where}
\]

\( M_t = \frac{H_t + L_t + C_t}{3} \), respectively: maximum price, minimum price and closing price;
\( SM_t \) – simple moving average;
\( D_t = \frac{\sum_{i=1}^{n}[M_{t-i} + SM_t]}{n} \)

So, the exogenous neural network assumes using the ten indicators described above as an input, while the next period closing price is supposed to be used as an output (22).

\[
C_t = f(u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}), \text{ where}
\]

\( u_1 \) – 10-day moving average;
\( u_2 \) – 10-day weighted moving average;
\( u_3 \) – momentum;
\( u_4 \) – fast stochastic oscillator K%;
\( u_5 \) – relative strength index;
\( u_6 \) – slow stochastic oscillator D%;
\( u_7 \) – MACD;
\( u_8 \) – A/D indicator;
\( u_9 \) – percent range %R;
\( u_{10} \) – commodity channel index.
It is important to note that the functional dependence of the t time closing price from the indicators described above is presented for the past period.

The main task of training an artificial neural network is to select a function (vector) that can identify which class an object belongs to based on a set of its features. The error function, in this case, has the following form \( (23) \):\[
L(D, F) = \sum_i \left[ F(x_i) \neq y_i \right]
\]

Thus, a classification neural network model is proposed for predicting financial time series. The ten technical indicators described above are supposed to be used as input signals, and the closing price of the next period will serve as an output signal.

The variety of indicators helps to evaluate the financial series in all possible aspects: the strength and direction of the trend, the overbought or oversold of a financial asset, and the market sentiment.

Let us give an example of building an exogenous neural network consisting of 14 hidden neurons.

Below is a table for estimating forecast accuracy for the model variant under consideration.

**Table 1.** Results of the exogenous neural network with the number of hidden neurons equal to 14 (in %).

|   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|   | 58,11 | 58,22 | 58,28 | 58,26 | 58,34 | 58,37 | 58,45 | 58,46 | 58,34 | 58,27 |

The main estimates of modelling quality are presented in the table below.

**Table 2.** Quality evaluation of modelling a dynamic neural network with the number of hidden neurons equal to 14 (in %).

| Range | Minimum | Maximum | Average |
|-------|---------|---------|---------|
| 0.35  | 58,11   | 58,46   | 58,31   |

The exogenous neural network under consideration shows an average forecast accuracy of 58.31% and an accuracy range of 0.35. Figure 1. shows a graph for estimating the accuracy of the model.
4. Conclusion

Let us highlight the main advantages of ANNs in comparison with other methods and models of forecasting. ANNs are suitable for implementing tasks with indefinite process development patterns and the unknown relationship between input and output data. ANNs operate stably when processing a significant number of low-information, noisy input signals. ANNs have an adaptive capacity. ANNs can solve complex and resource-intensive tasks in a limited time frame. In their hardware implementation, ANNs appear to be potentially resistant to adverse conditions and show little if any loss of productivity.

References

[1] Hlavovská K and Koncz P 2012 Application of Artificial Intelligence and Data Mining Techniques to Financial Markets *Economic Studies & Analyses/Acta VSFS* 6(1)
[2] Tang Y, Ji J, Gao S, Dai H, Yu Y and Todo Y 2018 A pruning neural network model in credit classification analysis *Computational Intelligence and Neuroscience*
[3] Pyo S, Lee J, Cha M and Jang H 2017 Predictability of machine learning techniques to forecast the trends of market index prices: Hypothesis testing for the Korean stock markets. *PloS one* 12(11) e0188107
[4] Chae Y T, Horesh R, Hwang Y and Lee Y M 2016 Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings *Energy and Buildings* 111 184-194
[5] Ivanyuk V and Soloviev V 2019 Efficiency of Neural Networks in Forecasting Problems. In *2019 Twelfth International Conference" Management of large-scale system development"*(MLSD) 1-4
[6] Khashei M and Bijari M 2010 An artificial neural network (p, d, q) model for timeseries forecasting *Expert Systems with applications* 37(1) 479-489
[7] Kim J, Moon J, Hwang E and Kang P 2019 Recurrent inception convolution neural network for multi short-term load forecasting *Energy and Buildings* 194 328-341
[8] Wang S, Zhang N, Wu L and Wang Y 2016 Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method *Renewable Energy* 94 629-636
[9] Zeng Y R, Zeng Y, Choi B and Wang L 2017 Multifactor-influenced energy consumption forecasting using enhanced back-propagation neural network *Energy* 127 381-396
[10] Elizarov M, Ivanyuk V, Soloviev V and Tsvirkun A 2017 Identification of high-frequency traders using fuzzy logic methods In *2017 Tenth International Conference Management of Large-Scale System Development*(MLSD) 1-4
[11] Soloviev V 2018 Fintech ecosystem in Russia. In *2018 Eleventh International Conference" Management of large-scale system development"*(MLSD) 1-5
[12] Wang L, Zeng Y and Chen T 2015 Back propagation neural network with adaptive differential evolution algorithm for time series forecasting *Expert Systems with Applications* 42(2) 855-863
[13] Qiu M, Song Y and Akagi F 2016 Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market *Chaos, Solitons & Fractals* 85 1-7
[14] Taheri K, Hasanipanah M, Golzar S. B and Abd Majid M Z 2017 A hybrid artificial bee colony algorithm-artificial neural network for calculating the blast-produced ground vibration. *Engineering with Computers* 33(3) 689-700
[15] Ivanyuk V 2018 Econometric Forecasting Models Based on Forecast Combination Methods. In *2018 Eleventh International Conference" Management of large-scale system development"*(MLSD) 1-4
[16] Koroteev M V, Terelyanskii P V and Ivanyuk V A 2016 Arithmetic of fuzzy numbers in generalized trapezoidal form *Journal of Mathematical Sciences* 216(5) 696-701
[17] Koroteev M V, Terelyanskii P V and Ivanyuk V A 2016 Fuzzy Inference as a Generalization of the Bayesian Inference *Journal of mathematical sciences* 216(5) 685-691