Investigation of financial market prediction by recurrent neural network

Nijolė Maknickienė[1] [a], Aleksandras Vytautas Rutkauskas[1], Algirdas Maknickas[2]
[1] Faculty of Business Management, Vilnius Gediminas Technical University, Saulėtekio al. 11, Vilnius, Lithuania
[2] Faculty of Fundamental Sciences, Vilnius Gediminas Technical University, Saulėtekio al. 11, Vilnius, Lithuania

Received 17 August 2011, accepted 17 October 2011

Abstract. Recurrent neural networks as fundamentally different neural network from feed-forward architectures was investigated for modelling of non linear behaviour of financial markets. Recurrent neural networks could be configured with the correct choice of parameters such as the number of neurons, the number of epochs, the amount of data and their relationship with the training data for predictions of financial markets. By exploring of learning and forecasting of the recurrent neural networks is observed the same effect: better learning, which often is described by the root mean square error does not guarantee a better prediction. There are such a recurrent neural networks settings where the best results of non linear time series forecasting could be obtained. New method of orthogonal input data was proposed, which improve process of *EVOLINO* RNN learning and forecasting.

Citations: Nijolė Maknickienė, Aleksandras Vytautas Rutkauskas, Algirdas Maknickas. Investigation of Financial Market Prediction by Recurrent Neural Network – *Innovative Infotechnologies for Science, Business and Education,* ISSN 2029-1035 – **2(11)** 2011 – Pp. 3-8.

Keywords: Financial Forecasting and Simulation; Time-Series Models; Dynamic Quantile Regressions; Dynamic Treatment Models; Forecasting Models; Simulation Methods; Neural Networks and Related Topics; Recurrent Neural Networks; *EVOLINO* learning algorithm; Non linear time series; Orthogonal inputs; Prediction of financial markets.
JEL: G17; C32; C53; C45.
Short title: Investigation of financial market prediction by RNN.

Introduction

Modeling of non linear processes is actual in two aspects. Realistic models of their history helps to understand the inner structure of nonlinearity. On the other hand, correct understanding of nonlinearity improves the prediction of these processes. For this purpose researchers use the most common Mackey-Glass system [1] as standards tester of non linear processes. The best recognized tools for the finance currency markets is in the last decade neural networks [2-4] or by Reinforcement-Learning Agents [5-6]. Intensive researches of financial market data by neural networks shows that the best learning stage of neural networks does not always lead to correct forecasting.

Financial markets could be explained by using profitability, reliability or risk investment models and analysis methods. Opposite to statistical analysis there could be sophisticated reinforcement learning agents [6] or neural networks [2-4]. The best suited neural networks for the recursive nonlinearity are *Recurrent Neural Networks* (RNN). Behavior of time series in financial, stock or currency markets are influenced by psychology of trades and are strongly non linear and hardly predictable. Using the RNN in modeling of financial time series is based on founding an acceptable learning model for RNN's. RNN's are fundamentally different from feedforward architectures in the sense that they not only operate on an input space but also on an internal state space [7-8]. For the better improvement of RNN learning the *EVOLINO* algorithm [1] could be selected because it very clearly shows training and validation of the recurrent neural network for non linear data inputs.

The goal of present works is to understand how RNN works for modeling and prediction of the financial markets, their behavioral analysis and paying attention to the acceptance of the chosen method for the anticipation. Author of Ref. [9] was exploring the human mind distinguished reproductive thinking, which only echoes the familiar issues, and productive thinking, that creates something new. In order to solve a specific task certain knowledge is needed.

Meanwhile, not everybody, who has the knowledge required for the task, is able to use it productively. There is no direct link between the past experience and new (productive) thinking. By observing the RNN learning and forecasting the same effect is observed - better learning, which is described by the *Root Mean Square Error* (RMSE), does not guarantee a better prediction.

The aims of this paper are to find the best conditions, where *EVOLINO* RNN becomes a good instrument of financial markets prediction. It will be investigated collection of RNN parameters like number of epochs, number of neurons to achieve strong learning of RNN and good prediction of financial markets. The input orthogonalization method is proposed attaining this goal.

[a] Corresponding author, cell: +370 (687) 25206; email: *nijole.maknickiene@vgtu.lt*

1. *EVOLINO* Learning Algorithm. Description

Neural networks aid to monitor of the non linear processes in the learning activity. The comparison of various methods evaluates neural networks learning algorithms of non linear processes and increase their prediction accuracy. Schmidhuber et al. [10] introduced a general framework of sequence learning algorithm *EVOlution of recurrent systems with LINear Outputs (EVOLINO)* [1]. *EVOLINO* uses evolution to discover good RNN hidden node weights, while using the methods such as linear regression or quadratic programming to compute optimal linear mappings from the hidden state to the output.

When quadratic programming is used to maximize the margin, it is impossible to obtain the first evolutionary recurrent support vector machines. *EVOLINO*-based Long Short-Term Memory (LSTM) can solve tasks that Echo State nets cannot [1]. There was introduced a new class of recurrent, truly sequential SVM-like devices with internal adaptive states, trained by a novel method called *EVOlution of systems with KErnel-based outputs (EVOKE)*, an instance of the recent *EVOLINO* class of methods.

EVOKE evolves recurrent neural networks to detect and represent temporal dependencies while using quadratic programming/support vector regression and pseudo-inverse regression. *EVOKE* is the first SVM-based mechanism which knows how to classify a context-sensitive language. It also outperforms recent state-of-the-art gradient-based *Recurrent Neural Networks* (RNNs) on various time series prediction tasks. RNN learning is used for context-sensitive languages recognition and is a difficult and often increasing problem for standard RNNs, because it requires unlimited memory resources.

For these array of problems investigated by authors of Ref. [1], [10-12], *EVOLINO* based LSTM learns in approximately 3 min on average and it is able to generalize substantially better that gradient-based LSTM. With *EVOLINO* it makes impossible to learn functions composed of multiple superimposed oscillators such as double sine and triple sine. Investigated network reached good learning and still makes very accurate predictions [1], [10-12]. The Mackey-Glass system is a standard benchmark for non linear time series prediction. Authors of Ref. [1] show deviation between the curves of *EVOLINO* generated and Mackey-glass system. *EVOLINO* is capable of making precise (0.0019) prediction in tasks like the Mackey-Glass benchmark.

The block diagram of *EVOLINO* recurrent neural network is shown in Fig 1. *EVOLINO* RNN forms LSTM network with $N = 4n$ memory cells, where N is total amount of neurons and n is amount of memory cells. The genetic evolution algorithm is applied to each quartet of memory cells separately. The cell has an internal state S together with a forget gate (G_F) that determines how much the state is attenuated at each time step. The input gate (G_I) controls access to the cell by the external inputs that are summed into the Σ unit, and the output gate (G_O) controls when and how much the cell fires.

Dark blue nodes represent the multiplication function and the linear regression Moore-Penrose pseudo-inverse method used to compute the output (light blue circle) [13]. The detail description of *EVOLINO* RNN algorithm could be found in Ref. [1], [10].

2. Inputs of Recurrent Neural Network

Authors of Ref. [14] analyzed what inputs should be chosen to receive the best models and got the inputs allowing a better prediction. They have found, that it is common to use the orthogonal inputs, where orthogonality of inputs is equivalent to orthogonality of n dimension vectors. The orthogonality of vectors is result of following inner product for two vectors f and g:

$$\langle f, g \rangle_w = \sum_n f(n) * g(n) * w(n) \tag{1}$$

where $w(n)$ is a non-negative weight vector of the definition of inner product. These vectors are orthogonal if above described inner product is zero:

$$\sum_n f(n) * g(n) * w(n) = 0 \tag{2}$$

This method was selected for finding the best inputs in *EVOLINO* learning process too. The most tools of financial prediction are used for searching dependences between time series of financial indicators or similar patterns in these time series. For this purpose the time series orthogonalization were exploit as follows

$$\mid \sum_n f(n) * g(n) \mid = \varepsilon. \tag{3}$$

where absolute value of scalar multiplication of vectors ε describes degree of orthogonality, because true orthogonality (2) could not be reached for time series of financial markets and non-negative weight vector is $w(n) = 1$. Prediction of one time series output were obtained by the two most orthogonal time series inputs. The influence of data orthogonality were investigated in the range of $\varepsilon \in [0.00001 \div 0.001]$.

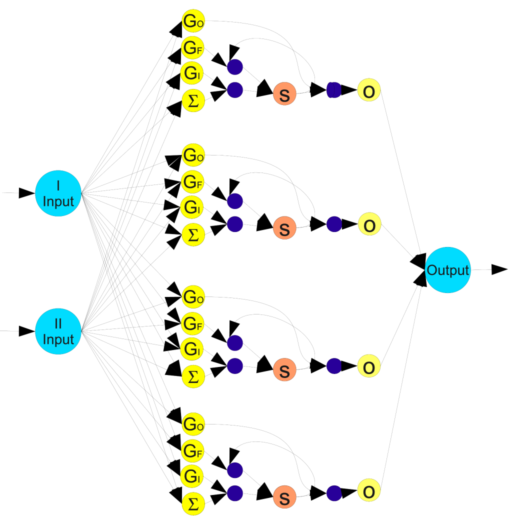

Fig 1. LSTM network with four memory cells.

3. Reproductive and productive learning

Behaviour of human brains could be divided into a productive and reproductive thinking. Reproductive thinking only echoes acquired knowledge and productive thinking creates something new. These brain processes are not straightforward. RNN algorithm ability to learn non linear process is measured and evaluated by RMSE aid. By studying non linear processes, such as a stock or currency markets or fluctuations in solar activity, or others, the RNN prediction is very important research in today areas. It could be distinguished into two major aspects of forecasting: i) how many correct data points of the process could be predicted; ii) how many correct directions of the process could be predicted. If we set itself the objective of accurately predicting of values of the non linear process, we will be facing the problem, as the average value of deviation is acceptable. For such processes, as the shares of a stock or currency market, the prediction of the direction is sufficient for the making the reasonable decision of the future investment. This work will attempt to prove or reject that. The proving or rejection will be confirmed in the investigation of RNN *EVOLINO* algorithm with finding of parameters where the prediction is the best.

4. Criteria learning and predictions

For investigation of *Root Mean Square Error* (RMSE) learning and prediction processes we used program framework [15] adopted for multiple inputs. We selected two parameters - RMSE and correlation for comparison of learned and predicted time series. RMSE is often used in RNN as a learning criteria. Learning of RNN means tests of trained neural network, where data for tests were used after wash out of 1/3 start data of training data. Observation of good RMSE result for learning some times do not imply good forecasting. For this purpose we use correlation coefficient too. As it was mentioned above, moving direction of stock shares or currency ratios is more important than the prediction of their exact values.

The correlation coefficient is located in range $[(-1) \div (+1)]$. A value of $(+1)$ implies that a linear equation describes the relationship between data ranges and predictive ranges perfectly, with all data points lying on a line in which Y increases as X increases. A value of (-1) implies that all data points lie on a line in which Y decreases as X increases.

5. Results and discussion

5.1. Selection of right numbers of epochs

Investigating some certain phenomena of artificial neural networks to work as memory structure or as a predictor is very important to clarify the behavior of the RNN.

Table 1. Dependence of test RMSE of learning from the number of epochs with data of *USD/JPY* and Gold

Number of epochs	RMSE of learned RNN
16	-0.008298
32	-0.007691
64	-0.007232
72	-0.006012
76	-0.002237
80	-0.002664
120	-0.002270
164	-0.001667
172	-0.001652
188	-0.001523
200	-0.001824
220	-0.001671

It is very important to have not only a well-selected input data and training ranges, but also a good selection of epochs and the number of neural network. Epochs number describes the number of times when NN data are processed, and may wrongly appear that the higher the number of epochs leads to the better learning and prediction. Finally, it was studied the RMSE dependence on the number of epochs, taking the familiar orthogonal data ranges of currency market *USD/JPY* with using of *XAU/USD* as a additional input for improving of convergence.

The obtained results of learning RMSE are taken in Table 1. Results shows that a small number of epochs does not provide RNN learning. Only 76 epochs starts to learn RNN. RMSE dependence on epochs shows that after reached of 164 epochs learning is stabilizing and further increasing of the epochs does not make sense.

5.2. The importance of the number of neurons

The variance of the number of neurons is very important neural networks parameter in the RNN learning. At first glance it may seem that the more neurons are used the better the prediction result will be. But a large amount of neurons took more calculation time and it is important too. Therefore, it is necessary to find optimal number of neurons, which are able to learn and predict data of time series. The results of dependence of RMSE and correlation coefficient from the number of neurons were obtained with of data 85 points of *USD/JPY* with additional input of *XAU/USD*.

The studying the dependences of the number of neurons on learning RMSE could be seen that learning slowly increases, when the number of neurons increases from 16 to 64 neurons. Starting from 68 neurons learning suddenly increases 26 times and the behavior of RNN learning becomes excellent. Investigation of the dependences the number of neurons on the RMSE and correlation coefficients of prediction could be shown in Table 2.

The three areas of distinct neural networks amount have been found.

Table 2. Dependence of RMSE and correlation coefficient from the number of neurons.

Number of neurons	RMSE of learning	RMSE of prediction	Correlation of prediction
16	-0.009911	1.207037	-0.275067
20	-0.010139	0.269538	-0.334100
24	-0.006452	0.132038	0.000400
28	-0.005201	0.658168	-0.193200
32	-0.004240	0.229054	0.053900
36	-0.003299	0.478200	0.011100
40	-0.002659	0.190520	0.035680
44	-0.002075	0.403500	0.184900
48	-0.001483	0.419292	-0.12925
52	-0.000941	0.258178	0.399926
56	-0.000354	0.363497	0.157300
60	-0.000110	0.560873	-0.042367
64	-0.005167	0.945073	-0.015375
68	-1.600e-29	0.627907	0.068900
72	-8.580e-30	0.757063	-0.051500
76	-1.350e-29	0.533008	-0.383733
80	-5.587e-30	0.312393	-0.224267
84	-7.252e-30	0.665393	-0.272200
88	-1.450e-30	0.481330	-0.227167
100	-2.300e-31	0.882620	-0.292600

The first area is for numbers of neurons from 16 to 40 where averages of correlation coefficients are in interval $[-1 \div 0.1]$. This proves that there are not enough neurons in RNN to learn and to predict. The second area for number of neurons is from 52 to 56 where averages of correlation coefficients are in interval $[0.1 \div 1]$, this proves that RNN try to learn and predict the data. All values of correlation coefficient in this area are in interval $[0 \div 1]$, this proves that RNN predict directions of *USD/JPY* very well.

Third area for numbers of neurons is from 60 to 100 where averages of correlation coefficients fitt interval $[(-1) \div 0.1]$, this proves that increasing of the number of neurons improve learning and RMSE of learning, but suddenly RNN stop to predict. Correlation coefficients versus amount of neurons are presented in Fig. 2. Presented curves shows, that a zone of amount of neurons exists where maximum correlation could be achieved.

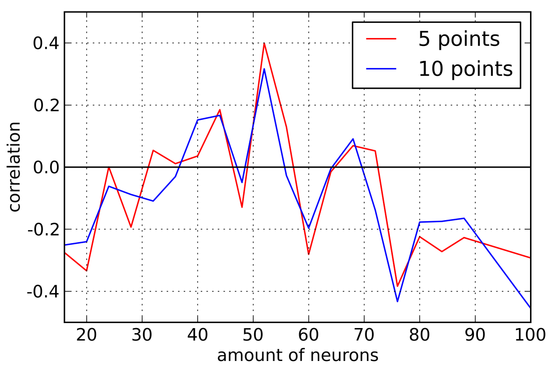

Fig 2. Correlation coefficients versus amount of neurons

The similar results are obtained for five and for ten points prediction.

5.3. Variation of data amount

The last stage of investigation was the variance of the input data size. It was important to know how many days are sufficient to monitor the financial or foreign exchange market in order to obtain reliable forecasting using RNN. In this purpose dependence the number of data on RMSE of learning, RMSE of predicting and correlation of prediction were obtained. Study of dependences the number of data to the RMSE and correlation coefficients and finding of suitable, predictable RNN shows that the RNN behave in the same way as in previous investigations. There are three distinct neural networks areas in the number of data: under learned, best learned and over learned. Dependence the number of data on learning and prediction RMSE and correlation coefficients for *USD/JPY* and *XAU/USD* inputs and *USD/JPY* output is presented in Table 3. Three kinds of behavior of learning and prediction are given in Fig. 3.

The first area could be separated in which there is not enough data for RNN learning and prediction. The second area of numbers of input data is the best area for RNN learning and prediction of input data. The third area showed that increasing of the number of data improves learning and RMSE of learning, but RNN stop to predict and the further increase of number of neurons do not imply better prediction.

All three studies have shown that the RNN prediction could be obtained when the neural network parameters such as epochs, number of neurons and the number of data are in a certain range.

Table 3. Dependence on learning and prediction RMSE and correlation coefficients from the number of data

Number of data	Number of neurons	RMSE of learning	RMSE of prediction	Correlation of prediction
50	36	-0.000054	0.09935	0.4528
57	36	-0.000415	0.29459	0.5328
65	36	-0.001570	0.2077	0.7571
70	36	-0.002385	0.29763	0.9103
76	36	-0.002808	0.07895	0.1006
85	36	-0.001714	0.1286	0.8726
85	64	-0.007561	0.09726	0.3360
90	64	-0.001452	0.1564	-0.0286
95	64	-0.000714	0.18421	-0.7302
100	64	-0.000990	0.14608	-0.7320
105	64	-0.001245	0.3206	0.6198
110	64	-0.001777	0.23879	0.7114
115	64	-0.002061	0.2093	-0.2290
120	64	-0.004170	0.25555	-0.5788
125	64	-0.003538	0.16155	-0.7229
130	64	-0.003216	0.14268	-0.8211

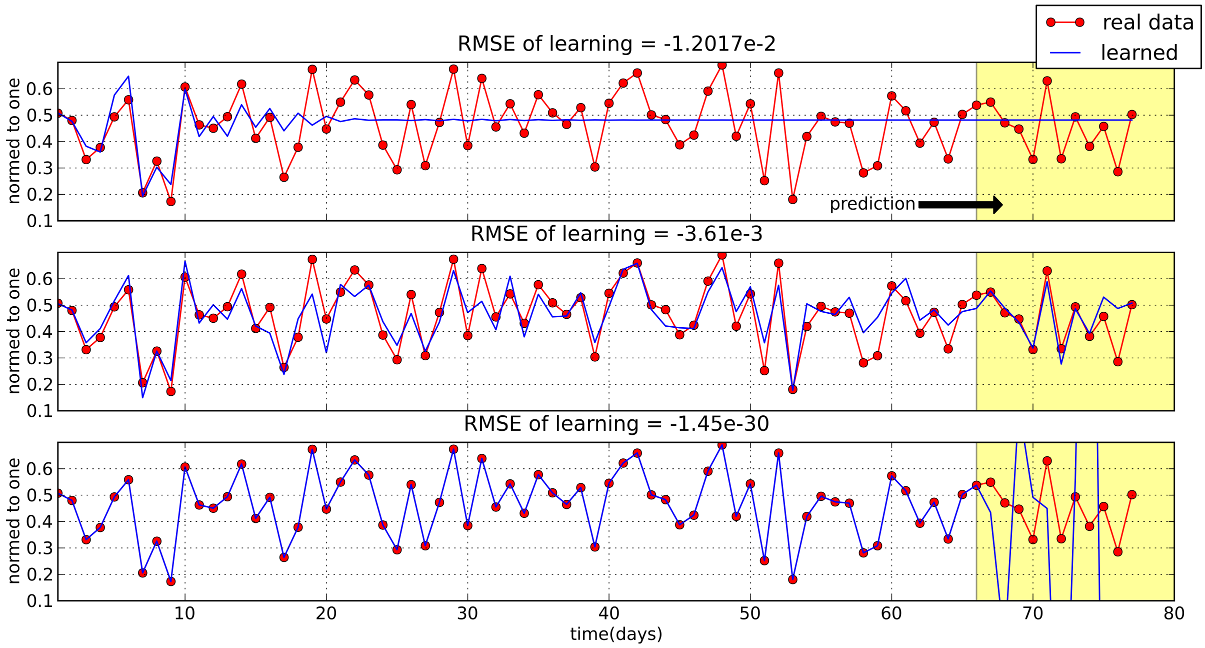

Fig 3. RNN learning and prediction. Three types of tests.

Our studies have shown the most 164 epochs is enough. Number of neurons must be in interval $[52 \div 56]$ and number of data in intervals $[80 \div 85]$ or $[105 \div 110]$. The averaged correlation coefficient of forecasting for this range of initial parameters of RNN were reached value of 0.32 and, in separate studies, even 0.9579.

6. Conclusions

The aim of the presented work was finding of the best conditions where the RNN makes the best prediction of currency markets. It was investigated that the prediction of the evolution of the *USD/JPY* exchange daily rates for 15 March 2010 in 5 following days. Data were collected from 1 Januar 2009 till 15 March 2010. *USD/JPY* exchange rates trained for the same period of *XAU/USD* data inputs. The obtained results show that.

1. The lowest values of orthogonality degree description coefficient ε improve stability of RNN learning and prediction of investigated non linear time series. The confirmation of quantitative dependences needs future investigation.
2. The learning and prediction of RNN, like the human brain productive and reproductive thinking, are independent and different. The better RMSE of learning do not guarantee the better achieving of prediction.
3. Combinations of parameters of RNN such as the number of epochs, data and neurons amount, determine different behavior of learning and prediction. weak learning without prediction; strong learning with prediction; excellent learning without prediction.
4. The investigation of financial data gives such group of parameters of RNN in *EVOLINO* algorithm, where RNN predict directions and values of a currency market. The group of RNN parameters for given data was found where the average of correlation coefficient of forecasting reaches maximum 0.938 and has value equal to 0.400.

References

1. J. Schmidhuber, D. Wierstra, and F. Gomez. Evolution: Hybrid neuroevolution / optimal linear search for sequence learning. – Proceedings of the 19th International Joint Conference on Artificial Intelligence (2005) 466–477.
2. M. Adya and F. Collopy. How effective are neural networks at forecasting and prediction. A review and evolution. – *Journal of Forecasting* 17 (1998) 481–495.
3. P. McNelis. Neural Networks in Finance: Gaining Predictive Edge in the Market. – Elsevier Academic press, 2005.
4. D. Plikynas. Explaining international investment patterns: a neural network approach. – *Information technology and control* 34(1) (2005) 42–50.
5. T. Ramanauskas and A. Rutkauskas. Empirical version of an artifical stock market model – *Monetary Studies* 1 (2009) 5–26.
6. A. Rutkauskas and T. Ramanauskas. Building an artificial stock market populated by reinforcement-learning agents. – *Journal of Business Economics and Management* 4 (2009) 329–341.
7. J. L. Elman. Finding structure in time. – *Cognitive Science* 14 (1990) 179–211.

8. M. Boden. A guide to recurrent neural networks and backpropagation. Technical report, DALLAS project. Report from the NUTEKsupported project AIS-8, SICS.Holst: Application of Data Analysis with Learning Systems, 2001.
9. E. Rimkutė. Mąstymas ir kalba (in lith.). – Vilnius: University press, 2007.
10. J. Schmidhuber, M. Gagliolo, D. Wierstra, and F. Gomez. *EVOLINO* for recurrent support vector machines. In: ESANN'2006 proceedings - European Symposium on Artificial Neural Networks, Apr. 2006, ISBN 2-930307-06-4 (2006) 593–598.
11. B. Schrauwen, D. Verstraeten, and J. Campenhout. An overview of reservoir computing: theory, applications and implementations. In: Symposium on Artificial Neural Networks, ISBN 2-930307-07-2. (2007) 471–482.
12. D. Wierstra, F. Gomez, and J. Schmidhuber. Modelling systems with internal state using *EVOLINO*. In: Conference on genetic and evolutionary computation GECCO. (2005) 1795–1802.
13. R. Penrose. A generalized inverse for matrices. – In: *Proceedings of the Cambridge Philosophy Society* 51 (1955) 406–413.
14. V. Rutkauskas, N. Maknickienė, and A. Maknickas. Approximation of dji, nasdaq and gold time series with *EVOLINO* neural networks. – 6th International Scientific Conference *Lithuania Business and Management*, May 13-14, 2010,Vilnius. 1 (2010) 170–175.
15. D. Goodman and R. Brette. Brian: a simulator for spiking neural networks in python. – *Frontiers in Neuroinformatics* 2(5) (2008).