FORECASTING ON FOREX MARKET WITH RBF AND PERCEPTRON NEURAL NETWORKS

[a]LUKÁŠ FALÁT, [b]ALEXANDRA KOTTILOVÁ

University of Žilina, Faculty of Management Science & Informatics, Univerzitná 8215/1, 010 26 Žilina
[a]Lukas.Falat@fri.uniza.sk, [b]Alexandra.Kottilova@fri.uniza.sk

Abstract: This work deals with an alternative approach in financial modelling - artificial neural networks approach. The aim of this paper is to show that this type of time series modelling is an excellent alternative to classical econometric modelling. At first, neural networks using methods of supervised machine learning are discussed. After explaining theoretical basis of ANN, these models are then applied to specific exchange rate (AUD/USD). Finally, the comparison between statistical models and RBF and perceptron neural networks is made to illustrate the sense of using ANN models.

Keywords: financial forecasting, neural networks, time series, RBF, perceptron,

1 Introduction to Financial Forecasting

Today, to be able to know a future value of some phenomena is a great advantage, regardless of an area of interest. There is no exception in economic world. Banks want to know interest rate of central banks, forex traders want to know how the value of EUR/USD will be tomorrow etc. Risk management departments of many financial institutions including banks and their financial analytics try to find a "guide" how to do this in the most efficient way in order to minimize their losses and maximize their profits.

The approach, which is used the most and which has been used for many years, is a statistical approach. This approach is represented mainly by ARIMA and GARCH models. However, it has been showed that this technique does not always provide sufficient results.

Because of that reason, other methods, using mainly the power of computers, have been created. Among these methods of machine learning, artificial neural networks, inspiriting by a human neural network, have become very popular. Today, these ANN models have become the interest of many prediction analyzers.

The first part of this article discusses basics and principles of neural networks. Then some basic principles of specific type of artificial neural networks – perceptron will be discussed. Perceptron is a very first representative of a supervised feed forward neural network. Later, various kinds of RBF networks as an upgraded version of supervised neural network will be depicted and then tested.

Later, models of ANNs will be applied to specific exchange rates using a programming application constructed by the author of this work.

Finally, ANN models of tested exchange rate will be compared with classical econometric models which were also quantified to make a reasonable conclusion whether it is worth using artificial neural networks in the economic predictions.

2 Fundamentals of Artificial Neural Networks

2.1 Artificial Neuron: Mathematical Background

The aim of mathematical neuron is a process identification. One tries to find an input-output function so that the output would have desired parameters and the predicted error would be minimal.

Let F: $x_t \in R^k \rightarrow y_t \in R^1$ is a projection assigning k-dimensional vector of inputs $x_t^T = (x_{1t}, x_{2t}, \ldots, x_{kt})$ one dimensional output y_t in specific time moment t.

Let G: $G(x_t, w_t): x_t \in R^k_{train} \rightarrow y_t \in R^1_{train}$ is a restriction of F. The task is then to find the values of w_t so that functional values of G would be so close to known sample as it is possible. Let E(w) is function defined as

$$E(w_t) = \sum_{x_t, y_t \in R^k_{train}} (G(x_t, w_t) - y_t) \tag{1}$$

Function defined by (1) represents squares of deviations of function G from expecting values of function F. If a minimum is found, G is adapted for approximation of F. Training or adaption is performed on training set. Validation set is used for validation of training network.

2.2 Perceptron Neural Network

The neural network called perceptron is a feed-forward type of network. It was one of the first neural networks constructed but due to its good properties it had been used till nowadays. It is one of the simplest types of supervised artificial neural networks. As it is feed forward neural network it contains only forward relations which can be realised only from lower layers to higher. The architecture is shown on figure 1.

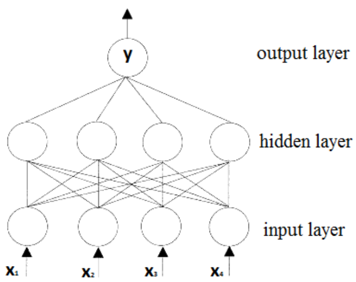

Figure 1 Perceptron neural network with 4 inputs

Vector $x_t^T = (x_{1t}, x_{2t}, \ldots, x_{nt})$ is a representative of input set of data. Output layer contains only one neuron and is represented by known given output. Between these layers thre is also a hidden layer containing hidden neurons. Using $G(x_t, w_t)$, the inputs are transformed into output when going through the network. The produced output is then compared to the historical output of a data set.

The network tries to get as close as possible to the given output by adapting its weights between input and hidden neurons and between hidden neurons and the output. The aim of learning is to have a trained network (and weights) so that G would approach F the most. The learning of perceptron, is based on back-propagation algorithm.

One should remember that the network quantifies its parameters only using training set. Afterwards a validation using validation set is performed. To compare predictive power of models numerical characteristics such as MSE or MAPE are used.

2.3 RBF Neural Network

Just like perceptron, RBF is a feed-forward neural network. The architecture is almost the same as of perceptron. The three biggest differences comparing to perceptron are calculation of potentials of hidden neurons, different activation function of processing (hidden) neurons and different activation function of output neuron.

In perceptron potentials of hidden neurons was just a simple linear operation of scalar product of vectors W and X. In RBF potential of j^{th} hidden neuron is defined as a difference of Euclidean distance given by vectors

$$u^j = \| x - w^j \|^2 \text{ , for j = 1,2,....s} \tag{2}$$

Moreover, RBF uses different types of activation functions (Gaussian or radial basis function) defined for j[th] hidden neuron as

$$\psi_2(u^j) = e^{\frac{-u^j}{2\sigma_j^2}} = e^{\frac{\|x-w^j\|^2}{2\sigma_j^2}} \text{ , for j = 1,2,....s} \tag{3}$$

where σ_j^2 is a variance of j[th] neuron.

Activation function of output neuron is also different, output neuron is always activated by a linear function y = x.

Learning process of RBF

RBF learning is performed by back-propagation algorithm defined as:

1. Introduction of input vector $x_t = (x_{1t}, x_{2t}, ..., x_{kt})$ from training set. X is then weighted by vector $w_t^j = (w_{1t}^j, w_{2t}^j, ..., w_{st}^j)$ and enters the hidden neurons. The output of every hidden neuron is counted as $o_t^j = \psi_2(u_t^j)$. These outputs are then weighted by vector $v_t^j = (v_t^j, v_t^j, ..., v_t^j)$. The total output of the network is then counted as $\hat{y} = \psi_3(u_t) = \sum_{j=1}^{s} v_t^j o_t^j$ This output is then compared to expected (known) value y and the error $e_t = y_t - \hat{y}_t$ is quantified.

2. On the base of back propagation of error e_t weights v are adapted

$$v_t^j \leftarrow v_t^j + \Delta v_t^j \tag{5}$$

where $\Delta v_t^j = \eta e_t o_t^j$ is an error term (for j = 1,2,…s)

3. Analogical adaptation of w_{rt}^j and variances σ_j^2 on the base of error term e_t using gradient method

$$w_{rt}^j \leftarrow w_t^j + \eta \Delta w_{rt}^j \tag{6}$$

$$\sigma_j \leftarrow \sigma_j + \eta \Delta \sigma_j \tag{7}$$

assuming that error terms Δ have the form

$$\Delta w_{rt}^j = e_t v_t^j \frac{1}{\sigma_j^2} \exp[-\sum_{r=1}^{k} \frac{(x_{rt} - w_{rt}^j)^2}{2\sigma_j^2}](x_{rt} - w_{rt}^j) \tag{8}$$

$$\Delta \sigma_j = e_t v_t^j \exp[-\sum_{r=1}^{k} \frac{(x_{rt} - w_{rt}^j)^2}{2\sigma_j^2}] \frac{(x_{rt} - w_{rt}^j)^2}{\sigma_j^3} \text{ ,}$$

j = 1,2,....s , r = 1,2,....k (9)

Iterative performing of epochs (repeating for every inputs of the training set) while the networks is considered to be learned (adapted) – error function achieves a minimum.

3 Application of Feed Forward Neural Networks to AUD/USD Exchange Rate

Exchange rate of AUD/USD will be applied to neural network model to show the predictive power of ANNs. The number of daily observations is 1044 and data are taken from 03/01/2007 to 03/01/2011. Due to finding out the predictive power of ANNs, data were divided into two parts – the training set contained 1002 values (03/01/2007 – 12/31/2010) and the validation set (for model verification) contained 42 observed values (01/01/2011– 03/01/2011). The validation set will be used for making ex-post (false) prediction to find out how the network manages with unknown data. The ANN modeling was performed by self-constructed application and econometric models were quantified in Eviews.

3.1 Prediction Process with Neural Networks

In order to have inputs entering into neural network we used Box-Jenkins model to determine autoregressive order in our data. By using Box-Jenkins identification procedure it was founded out that AUD/USD for the specific time period can be econometrically modeled using AR(0) process with only GARCH effect. Therefore, to have any inputs, the previous value (the first lag) was integrated into the input layer of this neural network.

Data from training set enter into network, then they are propagated through it and finally the value is produced by ANN. This value is then compared with a historical value (in the training part) and the error (MSE or RMSE) is counted. Adaptation of weights is performed using error values. At first, perceptron network was tested and the results are shown in table 1. The results are achieved after 2000 epochs.

Table 1 MSE_E achieved for AUD/USD (perceptron)

N / η	0,001	0,01	0,05	0,10	0,20	0,30	0,50
1	0,124056	0,087272	0,028243	0,014337	0,007650	0,005536	–
2	0,137236	0,064142	0,002228	0,002249	0,002344	0,002530	–
3	0,100385	0,002195	0,002262	0,002301	0,003671	0,003012	–
4	0,110228	**0,002173**	0,002259	0,002339	0,002663	0,003183	–
5	0,095508	0,002193	0,002274	0,003792	0,003343	0,003165	–
6	0,120620	0,002252	0,002397	0,002631	0,002950	0,003160	–
8	0,104851	0,003679	0,002395	0,002740	0,003202	0,003280	–
10	0,047155	0,002690	0,002828	0,003239	0,003357	0,003735	–
20	0,038187	0,003567	0,003240	0,003422	0,004718	–	–

The optimal speed of learning was set to η = 0,01, η = 0,05 respectively. Using this value of η parameter the network was able to well adapt on the specific data. The higher values were the reason for vibration of the network. The best value of MSE_E (0,002173) were achieved with 4 hidden values and with η = 0,01. However, very similar values were counted in some other cases. From that reason it can be concluded that perceptron trained by these data has the local minimum oscillating near the value of 0,002.

Next, RBF network as an upgrade of perceptron was tested. Extensions of this network were tested too - soft RBF, cloud RBF and granular RBF. Various RBFs were tested using only η = 0,05 as it was experimentally determined that this is a reasonable compromise between the expecting result and the time of adapting. Table 2 shows the tested results.

Table 2 MSE_E achieved for AUD/USD (RBF network)

Number of hidden neurons	Type of RBF			
	Classical RBF	Soft RBF	Cloud RBF	Granular RBF
1	0,0001836	0,0001834	0,0000813	0,0001842
2	0,0000841	0,0001834	0,0003693	0,0000659
3	0,0000932	0,0001834	0,0001348	0,0000523
4	0,0001285	0,0001834	0,0001514	0,0000368
5	0,0001685	0,0001834	0,0001255	0,0000332
6	0,0002609	0,0001834	0,0004750	0,0000407
7	0,0002416	0,0001834	0,0005516	**0,0000282**
8	0,0002714	0,0001834	0,0004215	0,0000341
10	0,0003209	0,0001834	0,0001889	0,0000324

As for RBF, the best neural network model for making ex-post predictions is the granular RBF network with architecture (1 – 7

-1). The RMSE for approximation was quantified to be 0,031223 and RMSE for ex-post predictions was quantified to be 0,005316. Figure two and three illustrate the graphic development of original and estimated / predicted values.

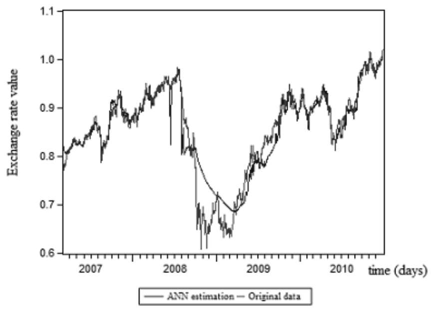

Figure 2 Original and estimated values of training set for AUD/USD

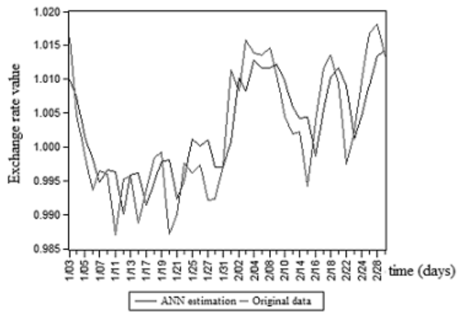

Figure 3 Original and estimated values of validation set of AUD/USD

Finally these neural network models quantified above were compared to traditional econometrical models which are the major tool for making predictions in various financial institutions. Comparison for the exchange rate AUD/USD is in the table 4. The models are compared according to their RMSE error in validation set.

Table 3 Comparison of Statistical and Neural Networks (AUD/USD)

	Model Type	**RMSE$_E$**
1.	Granular RBF (1 – 7 – 1)	0,005316
2.	AR(0) + TGARCH(1,1,1)	0,006205
3.	RBF (1 -12 – 1)	0,009063
4.	Perceptron (1 – 4 – 1)	0,046615

4 Conclusion

This paper dealt with neural networks forecasting and modelling. These models were applied to exchange rate modeling to compare their predictive power to traditional statistical models and methods. As for ANN, the parameters tested the most were the speed of learning and the number of hidden neurons. These parameters were tuned in order maximize the accuracy of predictions of the network.

As for neural networks, two feed forward neural networks were tested - perceptron and RBF network. As assumed from theory, perceptron did not achieve comparable results with RBF network. The value of 0,002 is a lot of worse than the error of RBF or statistical model. RBF network as an upgrade of perceptron was tested too. Moreover, extensions of this network were tested too – soft RBF, cloud RBF and granular RBF network. Generally speaking, RBF extensions provided better results than standard RBF. (granular RBF had the best results). The neural network models were finally compared to statistical models by using error function MSE and RMSE. From this statistics it can be concluded that neural network models (mainly RBF and its extensions) are comparable to statistical models, in some cases (extensions of RBF) ANN models even overcame statistical models. Therefore artificial neural networks as a representative of modern computing techniques for prediction process can be definitely considered as a great alternative to statistical model

Literature:

1. BOX , G. E. P. and JENKINS, G. M.: *Time Series Analysis, Forecasting and Control*. San Francisco, CA: Holden-Day, 1970
2. MARČEK, D.: *Some Intelligent Approaches to Stock Price Modelling and Forecasting*, Journal of Information, Control and Management Systems, Vol. 2, 2004
3. MARČEK, D., MARČEK, M.: *Artificial Neural Networks and Applications*, EDIS, University of Žilina, Žilina, Slovakia, 2006
4. MARČEK, D., MARČEK, M., PANČÍKOVÁ, L.: *Econometrics and Soft Computing*, EDIS, University of Žilina, Žilina, Slovakia, 2008
5. MARČEK, M.: *Multiple Statistical Analysis of Data and Time Series Modeling in Economy*, Silesian University, Opava, the Czech Republic, 2009
6. MONTGOMERY, D.C., JENNINGS, C.L., KULAHCI, M.: *Introduction to Time Series Analysis and Forecasting*, John Wiley & Sons, Inc, New Jersey, 2008
7. LEK, S., GUÉGAN, J.F.: *Artificial neural networks as a tool in ecological modeling*, Ecological Modelling 120 (1999) 65–73
8. ZHANG, G., PATUWO, B.E., HU, M.Y.: *Forecasting with Artificial Neural Networks: The State of Art*, International Journal of Forecasting 14, 1998. p35-p62

Primary Paper Section: I

Secondary Paper Section: IN