Proceedings of the International Conference on Information Technologies (InfoTech-2018) 20-21 September 2018, Bulgaria

DUAL MLP PAIRS WITH HIDDEN LAYER SHARING[1]

Ivan Blagoev, Janeta Sevova, Kolyu Kolev

Institute of Information and Communication Technologies
Bulgarian Academy of Sciences
e-mails: i.blagoev@iit.bas.bg
Bulgaria

Abstract: Artificial neural networks (ANN) are well known and widely used from decades. One of the most popular is the multilayer perceptron (MLP). MPL's general characteristic is that it has more than one layer. The most used topology of MLP has three layers (input, hidden and output). Layers are fully-connected, between each other (the input with the hidden layer and the hidden layer with the output). The general disadvantage of this topology is the lack of recurrent connections. The common usage of MLP is to solve only a single task. This research addresses these two common features of the MLP.
Key words: artificial neural networks, multilayer perceptron, topologies.

1. INTRODUCTION

A general description of classical ANN is a weighted directed graph [1]. Networks usually consist of neurons and connections between them. The way in which neurons are connected between each other is called topology of the network [2]. There are many topologies, but MLP is one of the most popular [3]. As its name implies, MLP consists of layers. Many layers can be used in MLP, but the simplest implementation is the three layers topology [4]. Neurons in a neighboring layers are usually fully connected [5]. It means that each neuron in one layer has connections to all neurons of the neighboring layer. The neurons in the input layer are connected with the external for the ANN world [6]. If the input has undesired noise, Kalman filter could be applied [7]. The output neurons are connected again with the external for the ANN world [6]. In the classical MLP networks the information goes from the input to the output and in some cases permutation of the neurons in the hidden layer

[1] This work was supported by private funding of Velbazhd Software LLC.

can be applied [8]. Such networks do not have feedback connections. Sometimes information can travel backward depending on the training algorithm and neurons activation function [9]. Such case is the backpropagation training where ANN total error is distributed backward in order ANN weights to be adjusted. In [10] a hierarchical topology of ANN is proposed that is suitable for big data analysis of time series. If MLP is accepted as monolithic base more advanced topologies can be created.

In this study, MLPs sharing use of the hidden layer is proposed. Such hidden layer sharing allows knowledge transfer between MPLs. Pairing of MLPs is also proposed thus recurrent connections can be used and effect of keeping of ANN memory can be achieved.

The paper is organized as follows: Section 1 introduces the problem; Section 2 presents an ANN topology proposition; Section 3 gives some experiment details; Section 4 concludes and some further ideas for research are pointed.

2. TOPOLOGY PROPOSITION

A common disadvantage of the classical MLP is the absence of backward links. Due to the lack of such links there is no short term memory presented in such ANN topology.

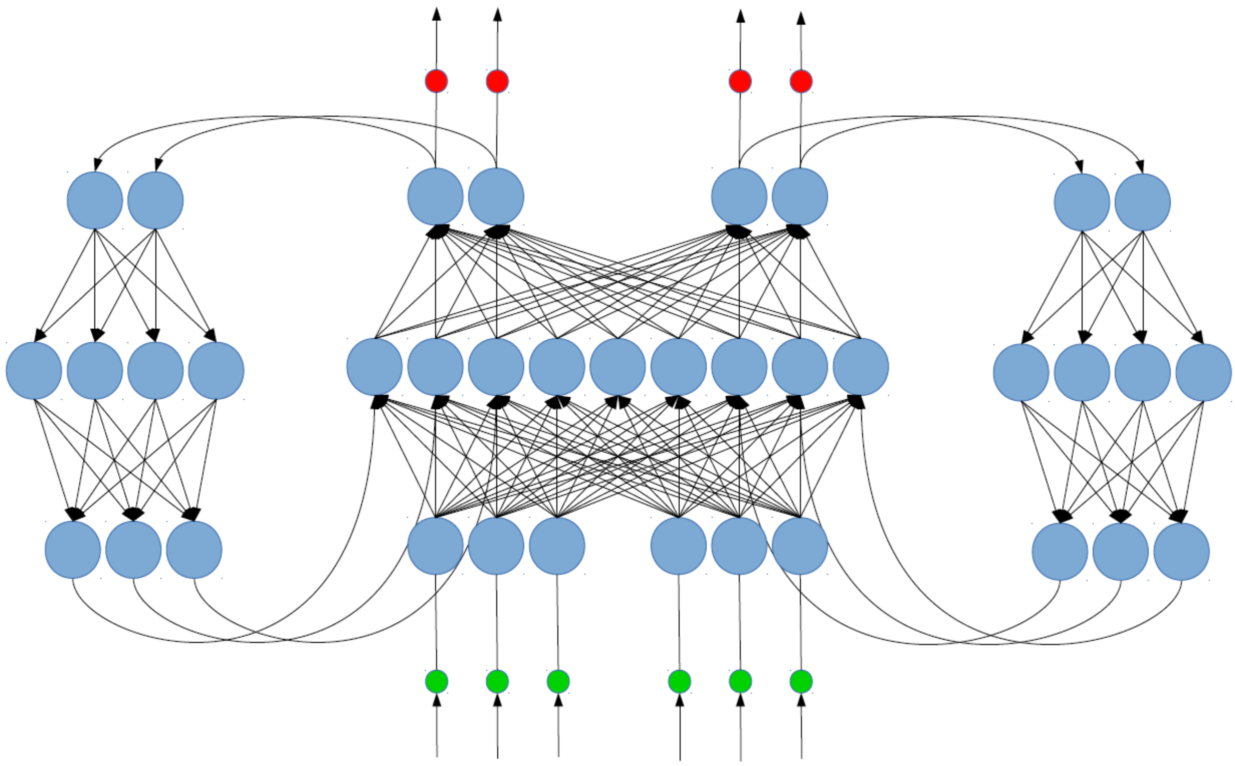

Fig. 1. MLP pairs with hidden layer sharing.

If simple MLPs are taken as building blocks of a more complex ANN, a variation of short term memory can be implemented (Fig. 1). The model proposed in this study uses four simple MLPs. The hidden layers of the two major MPLs are merged into a bigger shared layer. Such modification of the topology gives a way in which the information from the left MPL can travel to the right MPL and vice versa. Two other supportive MPLs are connected to the output of the major MPLs. The task of the supportive MPLs is to return some a part of the signal to the input of the ANN structure. The output of the supportive MPLs is supplied at the shared hidden layer. As expected values (target output) on the supportive MPLs the input for major MPLs is used.

3. EXPERIMENTS

The experiments are made with Encog Neural Networks Framework for Java.

Fig. 2. Currencies values for two months on daily basis - EUR/USD currency pair.

As input data, currencies values are used. One half of the ANN is supplied with EUR/USD values (Fig. 2). The other half of the ANN is supplied with USD/JPY values (Fig. 3). All data are normalized according to the activation function levels. Data are separated in three sets: training (75%), testing (20%) and validation (5%). Validation set appears to be the most important part of the training, to evaluate the forecasting capabilities of the ANN are evaluated.

It is clearly visible (Fig. 2 and Fig. 3) that the US economy is very related with the EU and Japanese economies. From forecasting point of view it is interesting that the correlation between both charts is negative. When EUR graph goes up the JPY graph goes down. It is something which ANN should learn and use effectively in the shared hidden layer.

Fig. 3. Currencies values for two months on daily basis - USD/JPY currency pair.

4. CONCLUSION

MLP based topologies can be very promising for financial time series forecasting. Two MLPs are modified in such way that they share common hidden layer. Both MPLs have the task to forecast currency values (EUR/USD for the first one and USD/JPY for the second one). The lack of recurrent connections is overcome by pairing each of the MLPs with another two MPLs. Pairs are completed in such a way that the output of one MPL is taken as input for its pairing MPL and the output of the pairing MPL is supplied to the original input. By such structure of four MPLs a knowledge transfer is achieved through the shared hidden layer and short term memory is presented by means of the pairs. As further research this approach can be combined with Generalized Nets [11, 12].

REFERENCES

[1] Balabanov, T., Genova, K. **(2016)**. Distributed System for Artificial Neural Networks Training Based on Mobile Devices. *Proceedings of the International Conference Automatics and Informatics.* Sofia, Bulgaria, Federation of the Scientific Engineering Unions John Atanasoff Society of Automatics and Informatics, ISSN 1313-1850, pp.49-52.

[2] Balabanov, T., Keremedchiev, D., Goranov, I. **(2016)**. Web Distributed Computing for Evolutionary Training of Artificial Neural Networks. *International Conference InfoTech, Varna - St. St. Constantine and Elena resort.* Bulgaria, Publishing House of Technical University - Sofia, ISSN 1314-1023, pp.210-216.

[3] Balabanov, T., Zankinski, I., Barova, M. **(2016)**. Strategy for Individuals Distribution by Incident Nodes Participation in Star Topology of Distributed Evolutionary Algorithms. *Cybernetics and Information Technologies, Institute of Information and Communication Technologies – BAS.* vol. 16, no. 1, ISSN 1311-9702, pp.80-88.

[4] Tomov, P., Monov, V. **(2016)**. Artificial Neural Networks and Differential Evolution Used for Time Series Forecasting in Distributed Environment. *Proceedings of the International Conference Automatics and Informatics.* Sofia, Bulgaria, ISSN 1313-1850, pp.129-132.

[5] Balabanov, T., Zankinski, I., Dobrinkova, N. **(2011)**. Time Series Prediction by Artificial Neural Networks and Differential Evolution in Distributed Environment. *Proceedings of the International Conference on Large-Scale Scientific Computing.* Sozopol, Bulgaria, Lecture Notes in Computer Science, Springer, vol. 7116, no. 1, ISBN 978-3-642-29842-4, pp.198-205.

[6] Keremedchiev, D., Barova, M., Tomov, P. **(2016)**. Mobile Application as Distributed Computing System for Artificial Neural Networks Training Used in Perfect Information Games. *Proceedings of the International Scientific Conference. UNITECH'16*, Gabrovo, Bulgaria, ISSN 1313-230X, pp.389-393.

[7] Alexandrov, A. **(2015)**. AD HOC Kalman filter based fusion algorithm for real-time Wireless Sensor Data Integration. *Proceedings of he Eleventh International Conference Flexible Quering Answering Systems.* ISBN 978-3-319-26153-9, pp.151-160, Warsaw, Poland.

[8] Zankinski, I., Stoilov, T. **(2016)**. Effect of the Neuron Permutation Problem on Training Artificial Neural Networks with Genetic Algorithms in Distributed Computing,.*Proceedings of the 24th International Symposium Management of Energy, Industrial and Environmental Systems.* ISSN 1313-2237, Bankya, Bulgaria, pp.53-55.
[9] Zankinski, I., Tomov, P., Balabanov, T. **(2017)**. Alternative Activation Function Derivative in Artificial Neural Networks,.*25th Symposium with International Participation - Control of Energy, Industrial and Ecological Systems.* Bankia, Bulgaria, John Atanasoff Union of Automation and Informatics, ISSN 1313-2237, pp.79-81.
[10] Atanasova, T., Barova, M., Balabanov, T. **(2016)**. Use of Neural Models for Analysis of Time Series in Big Data. *Publishing complex of "Vasil Levski" National Military University.* ISSN 1314-1937, pp.193-198.
[11] Tashev, T., Marinov, M., Monov, V., Tasheva, R. **(2016)**. Modeling of the MiMa-algorithm for crossbar switch by means of Generalized Nets. *Proceedings of the 2016 IEEE 8th International Conference on Intelligent Systems (IS).* Sofia, Bulgaria, ISBN 978-1-5090-1354-8, pp.593-598.
[12] Tashev, T., Monov, V. **(2012)**. Modeling of the hotspot load traffic for crossbar switch node by means of generalized nets. *Proceedings of the 6-th International IEEE Conference Intelligent Systems IS'12.* Sofia, Bulgaria, vol. 2, pp.187-191.