Hybrid Neo-Fuzzy Neural Networks Based on Self-Organization and Their Application for Forecasting in Financial Sphere

Yuriy Zaychenko[a] and Galib Hamidov[b]

[a] *National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Peremohy av. 37, Kyiv, 03056, Ukraine*
[b] *Information Technologies Department, Azershiq, str. K. Kazimzade 20, Baku, AZ1008, Azerbaijan*

Abstract
In this paper new class of deep learning – cascade neo-fuzzy neural network (CNFNN) was considered and investigated. Neo-fuzzy neuron with two inputs is used as a node of hybrid network. The experimental investigations were carried out during which the optimal parameters of neo-fuzzy network were found: number of inputs, training/test sample ratio and number of linguistic variables. Method GMDH was used for optimal structure construction of deep hybrid network. The experimental investigations of deep NFNN were carried out in the problem of market index forecasting at German stock exchange and Google share prices. The comparison experiments of the deep neo-fuzzy network with alternative methods GMDH and conventional cascade neo-fuzzy network were carried out and the efficiency of suggested hybrid NFN was estimated

Keywords1
Deep learning, GMDH, cascade neo-fuzzy network, parameters and structure optimization.

1. Introduction

Nowadays forecasting problems of share prices and market indicators attract great attention of investors and managers of invest funds. For forecasting at financial markets usually are applied methods of regressive analysis ARIMA, ARCH and GARCH methods, exponential smoothing [1].

But last years for solution of this problem fuzzy neural networks (FNN) were suggested [5-7]. Their main advantages are capability to work with fuzzy, incomplete and qualitive information and to utilize expert knowledge. Besides, FNNs have properties of high approximation due to FAT theorem [5] and interpretability. But for application of FNN in forecasting problems its necessary to train rule base and membership functions of fuzzy rules. This demands large computational resources and a lot of training time.

Last years new class of FNN- cascade neo-fuzzy neural networks (CNFNN) appeared [8], their main advantage is absence of necessity to train membership functions and only rule weights are to be trained using input sample. This enables to substantially cut training time and computational expense and to apply this class of FNN for high dimensional problems (Big Data).

Usually training of NN means the adjustment of weights between neurons. But efficiency of training can be substantially improved if to adapt not only neuron weights, but a network structure as well using training sample. For this aim the application of Group Method of Data Handling seems very promising. GMDH is based on the principle of self- organization and enables to construct structure of model in the process of its run. Besides, as building blocks for model simple sub-models called partial descriptions consisting of two variables are used [2-4]. That allows to GMDH to work with short training samples.

COLINS-2021: 5th International Conference on Computational Linguistics and Intelligent Systems, April 22–23, 2021, Kharkiv, Ukraine
EMAIL: zaychenkoyuri@gmail.com (Yu. Zaychenko); galib.hamidov@gmail.com (G. Hamidov)
ORCID: 0000-0001-9662-3269 (Yu. Zaychenko)

For the first time application of GMDH to construct structure of neural networks and to train its weights was suggested by A.G. Ivakhnenko and his collaborates (so-called "active neurons") [2-4]. In the next works method GMDH was successfully applied for construction of hybrid neuro-fuzzy networks with kernel activation functions [9], spiking neurons [10], wavelet functions [11,12] and other class of fuzzy neural networks [13]. But in these works, the important property of GMDH – application of basic models with only two inputs and small number of tunable parameters wasn't used. This property is very important for deep learning fuzzy networks and enable to cut number of adapted parameters and training time as well.

The goal of this paper is to find optimal parameters of deep CNFNN, construct its structure using GMDH, investigate its efficiency in the forecasting problem of share prices at stock exchange and to compare its efficiency with CNFNN of standard structure.

2. Experimental investigations of deep learning neo-fuzzy neural networks

The goal of experiments was to analyze the forecasting efficiency of deep learning neo-fuzzy neural network (NFNN) in the problem of forecasting share prices and market index of German stock exchange DAX, in particular find optimal parameters of neo-fuzzy network – number of inputs, number of fuzzy membership functions and optimal structure of deep NFNN.

As input data were taken average month values of market index DAX in the period since January 2010 to December 2016. Then total sample size was 80 elements – average month values. The input data is presented in the Table 1.

Table 1

Dynamics of German stock exchange index

Germany	2010	2011	2012	2013	2014	2015	2016
January	100	112,5032	96,8841	123,189	154,8426	140,4772	127,3783
February	91,32258	119,0592	107,396	122,1897	155,3302	148,9331	123,2948
March	96,74254	116,4645	109,996	122,595	154,2723	152,4478	131,1697
April	99,67204	124,9138	105,958	120,2768	156,6237	154,271	135,7509
May	89,38131	125,5138	98,0491	128,8783	159,2462	154,6962	135,1146
June	88,73958	123,1834	92,7264	127,6049	161,3321	150,7816	132,4843
July	92,74674	124,5796	96,2023	127,7148	157,7584	148,275	131,6502
August	94,39056	101,4666	103,017	132,616	147,5007	144,4888	141,021
September	97,24559	88,79165	111,930	135,6962	148,5193	134,0826	140,8114
October	106,9113	96,40158	112,958	143,4353	136,0243	136,9621	139,7606
November	109,9191	94,39673	111,048	147,9172	141,4704	140,8389	136,0866
December	110,4825	92,2427	118,689	151,2206	144,6012	138,8777	139,7837

Network training was performed by gradient descent method with adaptive steps and Widrow-Hoff method (6) in the sequential mode (see previous section). The goal of experiments was to find optimal parameters NFNN. In experiments the following parameters varied: inputs number (prehistory length), number of layers, number of linguistic variables (fuzzy sets/per variable), rules number and training/ test sample ratio N_{train}/N_{test} (%).

In the first experiment the number of layers was varied under different ratio training test sample and its influence on forecasting accuracy was explored. The corresponding results are presented in the Table 2. In denoting CNFNN (m,n,k) the first digit m indicates number of layers, the second digit n-inputs number, the third k-number of linguistic variables.

Table 2
Dependence forecasting accuracy versus number of network layers

Data points	Real	CNFNN(2,4,4)	CNFNN(3,4,4)	CNFNN(4,4,4)	CNFNN(5,4,4)
31	107.7221	130.2851	93,3602	110,1878	94,9563
32	110.0536	130.4976	93,0511	109,9537	94,6139
33	109.9942	131.1365	92,7585	109,7220	94,2873
34	107.4861	131.82	92,4814	109,4927	93,9759
35	105.812	132.1484	92,219	109,2658	93,6788
36	108.7387	132.4245	91,9706	109,0412	93,3955
37	109.2837	133.0117	91,7354	108,8190	93,1252
38	105.0767	133.6728	91,5126	108,5990	92,8674
39	102.8346	134.0803	91,3017	108,3813	92,6215
40	105.3798	134.4172	91,1021	108,1659	92,3870
MAPE, %	-----------	25.78	23,8	14,3	15,6

In the Figure 1 the corresponding dependence of criterion MAPE on layers number is presented.

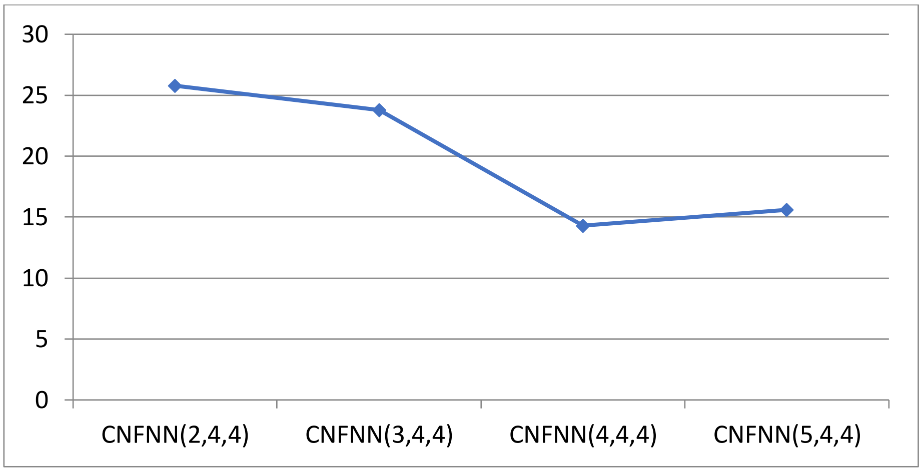

Figure 1: Dependence of MAPE (%) versus number of layers for ratio N_{train}/N_{test} = 50/50

As it follows from Figure 1 the optimal layers number for considered problem is equal to 4.
Further investigation of MAPE dependence on inputs number was carried out. The corresponding results are presented in Figure 2 for ratio N_{train}/N_{test} =50/50.

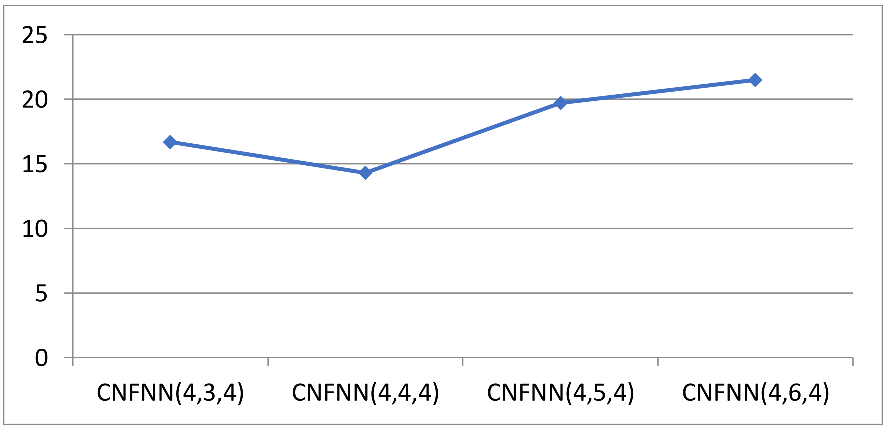

Figure 2: Dependence of criterion MAPE (%) on inputs number

As it follows from the presented results the optimal inputs number exists which, in general case depends on ratio N_{train}/N_{test}. The optimal inputs number for ratio $N_{train}/N_{test} =$ 50 equals to 4.

The important parameter for deep NFN is a number of linguistic variables (fuzzy sets per one variable). The corresponding investigations were carried out and the results are presented in the Figure 3.

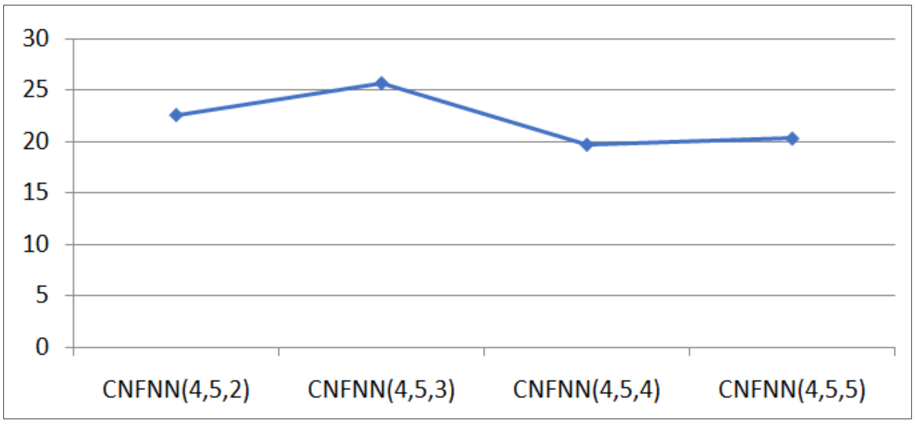

Figure 3: Dependence of criterion MAPE (%) on the number of linguistic variables

The presented results show optimal number of linguistic variables is equal to 4 for the considered problem.

The investigations of ratio N $_{train}$/N $_{test}$ dependence on forecasting accuracy were carried out.

The corresponding results of forecasting accuracy versus number of layers for different rations N_{train}/N $_{test}$ are presented in figure 6 and in the Table 3.

Table 3

Forecasting accuracy dependence on layers number

N_{train} /N_{test}	50-50	60-40	70-30	80-20	90-10
CNFNN(2,4,4)	25,78%	20,2%	11,052%	7,0012%	3,5213%
CNFNN(3,4,4)	23,8%	17,3%	10,5341%	5,9654%	3,2592%
CNFNN(4,4,4)	16,3%	15,4%	9,6584%	4,4325%	3,1952%
CNFNN(5,4,4)	19,6%	19,2%	12,9532%	6,3454%	3,2421%

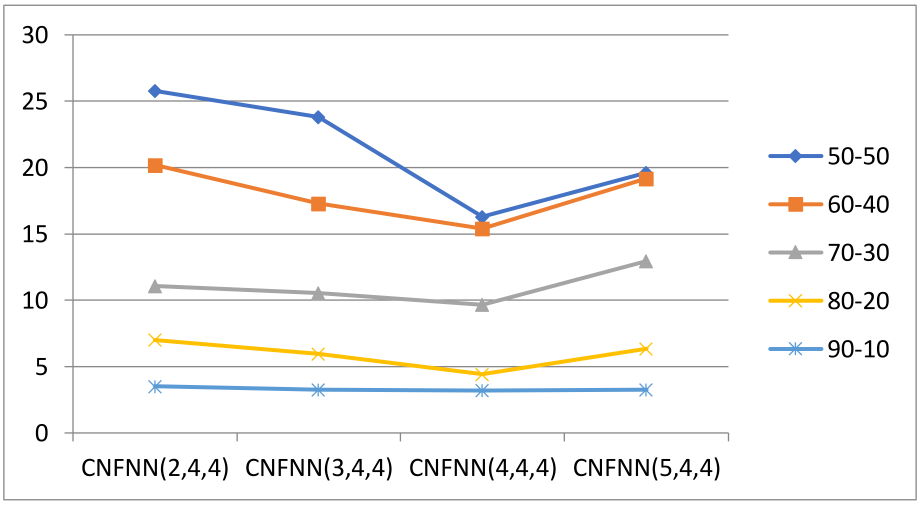

Figure 4: Dependence of MAPE (%) on layers number for different ratios N_{train} /N_{test}

The found dependence of criterion MAPE on inputs number for different ratios N $_{train}$/N $_{test}$ are presented in the Table 4.

Table 4
Forecasting accuracy MAPE (%) dependence on inputs number

N_{train} /N_{test}	50-50	60-40	70-30	80-20	90-10
NFNN(4,3,4)	17,7%	17,642%	10,5329%	4,8543%	4,5213%
NFNN(4,4,4)	15,5%	15,4%	9,6584%	4,4325%	3,1952%
NFNN(4,5,4)	19,7%	16,5922%	8,5811%	3,2151%	1,6819%
NFNN(4,6,4)	21,5%	19,6483%	8,6954%	4,9623%	1,7651%

The forecasting accuracy versus number of linguistic variables was also investigated for different ratios N_{train} /N_{test} and presented in Figure 5.

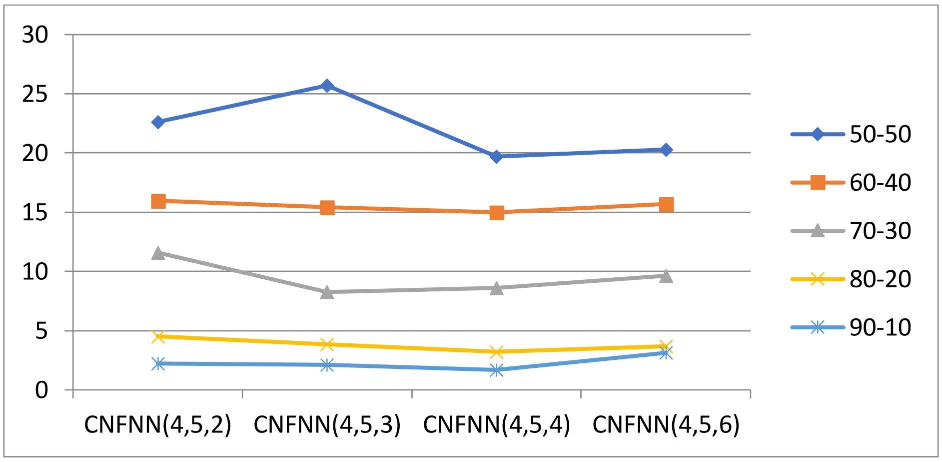

Figure 5: Dependence of MAPE (%) versus number of linguistic variables

As it follows from the presented results for each class of financial processes there exists optimal number of layers of NFNN. Under its further increase criterion MAPE on test sample begins to increase or stops to change. That is well comply with principle of self-organization of GMDH [2-4]. The similar dependence was detected for number of inputs and number of linguistic variables.

Further the similar forecasting experiments were carried out with FNN ANFIS. Number of inputs and number of linguistic values were taken 4. The efficiency comparison with the deep neo-fuzzy network with similar parameters values NFNN(4,4,4) was performed. The corresponding results for both networks are presented in the Table 5.

Table 5
Comparison of forecasting accuracy of NFNN and FNN ANFIS

MAPE	50-50	60-40	70-30	80-20	90-10
FNN ANFIS	19,7%	17,65%	12,54%	6,9554%	4,5614%
Deep NFNN(4,4,4)	15,5%	15,4%	9,6584%	4,4325%	3,1952%

As the results of comparison show the deep neo-fuzzy neural network has higher forecasting accuracy than ANFIS. Additional advantages of NFNN are less computational complexity and less training time due to lack of necessity to adjust membership functions These properties enable to use NFNN in Big Data forecasting problems.

In the process of investigations GMDH was applied to construction of optimal structure of hybrid cascade network. In this research Google shares close prices since August till December 2019 were forecasted. The process of hybrid network structure generation which was obtained by GMDH algorithm is presented in Figure 6 [14].

X1 A0 = f(X1, X2); B0 = f(A0, A4); C0 = f(B0, B3); D0 = f(C0, C1).
X2 A1 = f(X1, X3); B1 = f(A1, A2); C1 = f(B1, B2).
X3 A2 = f(X2, X4); B2 = f(A1, A5);
X4 A3 = f(X1, X5); B3 = f(A3, A4).
X5 A4 = f(X3, X4);
A5 = f(X2, X5).

Figure 6: The optimal structure of hybrid neo-fuzzy network

The optimal structure generated by GMDH was such: 6 neurons (A0, A1, A2, A3, A4, A5) at the first layer, 4 neurons (B0, B1, B2, B3) at the second layer, 2 neurons (C0,C1) at the third layer and one neuron (D0) at the last layer. All 5 inputs were used in the structure.

3. Conclusion

1. In this paper new class of deep learning networks-hybrid cascade neo-fuzzy neural network (NFNN) based on GMDH was developed and explored in the problem of forecasting market indicator of German stock exchange and Google share prices. In this type of deep networks neo-fuzzy neuron with two inputs is used as a node.
The experimental explorations were carried out during which the optimal parameters of hybrid neo-fuzzy network were determined.
2. The problem of optimal structure generation of hybrid cascade network was considered and for its solution method GMDH was applied and investigated.
3. The comparative experiments of the deep hybrid network with alternative methods GMDH and cascade network were carried out and forecasting efficiency of the suggested hybrid network was estimated and proved to be the best one.
4. After experiments it was detected the developed deep hybrid neo-fuzzy network is very promising for forecasting in the financial sphere. Besides, it's free from typical drawbacks of conventional deep learning networks.

4. References

[1] L. Lewis, Methods of forecasting economic indicators (transl. from English). Finance and Statistics, Moscow, 1986.
[2] A.G. Ivakhnenko, G.A. Ivakhnenko, J.A. Mueller, Self-organization of the neural networks with active neurons. Pattern Recognition and Image Analysis 4, 2 (1994): 177-188.
[3] A.G. Ivakhnenko, D. Wuensch, G.A. Ivakhnenko, Inductive sorting-out GMDH algorithms with polynomial complexity for active neurons of neural networks. Neural Networks 2 (1999): 1169-1173.
[4] G.A. Ivakhnenko, Self-organization of neuronet with active neurons for effects of nuclear test explosions forecasting. System Analysis Modeling Simulation 20 (1995): 107-116.
[5] M. Zgurovsky, Yu. Zaychenko, Fundamentals of computational intelligence- System approach. Springer, 2016.
[6] L.-X. Wang, J. M. Mendel, Fuzzy basis functions, universal approximation, and orthogonal least-squares learning. IEEE Trans. on Neural Networks 3, №5 (1992): 807-814.
[7] J.-S. Jang, ANFIS: Adaptive-network-based fuzzy inference systems. IEEE Trans. on Systems, Man, and Cybernetics. 23 (1993,): 665-685.
[8] T. Yamakawa, E. Uchino, T. Miki, H. Kusanagi, A neo-fuzzy neuron and its applications to system identification and prediction of the system behavior, in: Proceedings 2nd Intern. Conf. Fuzzy Logic and Neural Networks «LIZUKA-92». Lizuka, 1992, pp. 477-483.
[9] Ye. Bodyanskiy, N. Teslenko, P. Grimm, Hybrid evolving neural network using kernel activation functions, in: Proc. 17th Zittau East-West Fuzzy Colloquium, Zittau/Goerlitz, HS, 2010, pp. 39-46.

[10] Ye. Bodyanskiy, O.A .Vynokurova, A.I. Dolotov, Self-learning cascade spiking neural network for fuzzy clustering based on Group Method of Data Handling. J. of Automation and Information Sciences, 45, №3 (2013,): 23-33.
[11] Ye. Bodyanskiy, O. Vynokurova, A. Dolotov, O. Kharchenko, Wavelet-neuro-fuzzy network structure optimization using GMDH for the solving forecasting tasks, in: Proceedings 4th Int. Conf. on Inductive Modelling ICIM 2013, Kyiv, 2013, pp. 61-67.
[12] Ye. Bodyanskiy, O. Vynokurov, N. Teslenko, Cascade GMDH-wavelet-neuro-fuzzy network, in: Proceedings 4th Int. Workshop on Inductive Modeling «IWIM 2011», Kyiv, Ukraine, 2011, pp. 22-30.
[13] Ye. Bodyanskiy, O. Boiko, Yu. Zaychenko, G. Hamidov, Evolving Hybrid GMDH-Neuro-Fuzzy Network and Its Applications, in: Proceedings of the Intern conference SAIC 2018, Kiev, Ukraine, 2018.
[14] E. Bodyanskiy, Yu. Zaychenko, O. Boiko, G. Hamidov, A. Zelikman, The hybrid GMDH-neo-fuzzy neural network in forecasting problems in financial sphere, in: Proceedings of the International conference IEEE SAIC 2020, Kiev, Ukraine, 2020.