

Neural Network Associative Forecasting of Demand for Goods*

Vasiliy Osipov, Nataly Zhukova, Dmitriy Miloserdov

St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences, 39,
14 Line, St. Petersburg 199178, Russia

osipov_vasiliy@mail.ru, nazhukova@mail.ru, dmmil94@yandex.ru

Abstract. This article discusses the applicability of recurrent neural networks with controlled elements to the problem of forecasting market demand for goods on the four month horizon. Two variants of forecasting are considered. In the first variant, time series are used to train the neural network, including the real demand values, as well as pre-order values for 1, 2 and 3 months ahead. In the second variant, there is an iterative forecasting method. It predicts the demand for the next month at each step, and the training set is supplemented by the values predicted for the previous months. It is shown that the proposed methods can give a sufficiently high result. At the same time, the second approach demonstrates greater potential.

Keywords: Recurrent Neural Network, Machine Learning, Data Mining, Demand Forecasting.

1. Introduction

To date, the following methods are used to predict economic indicators: individual and collective expert assessments, forecast extrapolation, statistical methods, system-structural methods, methods of mathematical and computer modeling, neural network methods [1]. Hybrid approaches based on the analysis of economic time series in several stages are known [2, 3]. The temporal regularized matrix factorization approach is proposed for multivariate series [4].

Neural network methods are widely used [5-10]. In particular, the paper [8] noted the superiority of neural network algorithms over the ARIMA model in solving the problem of forecasting prices for agricultural products. The paper [9] investigates the applicability of artificial neural networks (ANNs) to economic analysis based on currency indicators. The article [10] proposes a hybrid neural network model and its application for stock price forecasting. Many works are devoted to forecasting the demand for various goods and services [11-18]. To predict the demand for utilities, such as water, heat and electricity, among others, neural networks of group method of data

* Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

handling (GMDH) [11], gray neural networks (GNN) [12], multilayer neural networks of direct distribution [13] and evolutionary algorithms [14] are used. Multilayer perceptron [15-17] and convolutional network [18] were used to predict the demand for consumer goods. However, they have certain drawbacks. In particular, GNN is characterized by a local optimum problem, which makes it impossible to achieve high accuracy of forecasts. Feedforward networks, such as a multilayer perceptron or convolutional network, do not have associative memory, which does not allow taking into account hidden patterns within the object under study.

Well-known architectures of recurrent neural networks (RNN) are also not without drawbacks. Their capabilities do not go far beyond classification and pattern recognition. The depth of information processing in them is low, at the same time there is a high redundancy of storage. There are problems with the extraction of previously processed information from memory, as well as with ensuring the stability of the RNS. Basically, the above shortcomings are related to the inadequacy of the structure and methods of associative-spatial information processing in such networks.

Architecture of RNN can be classified according to the criterion meet the requirements of working in real time. The first class includes perceptron-based RNN, such as recurrent multilayer perceptron, Elman network, and real-time RNN. These networks can work in real time, but do not differ in the depth of information processing. The second class of networks is designed for deeper data processing. These include Hopfield's associative memory, Cosco networks, and Kohonen's self-organizing maps.

To solve the existing problems, it is necessary to combine the properties of the two types of networks described above. Recurrent neural networks with controlled elements are proposed as such a tool. Such RNN are a structure of two identical layers of neurons and a control unit [19]. Each neuron of one layer is connected by synapses with all neurons of another layer, while synaptic connections between neurons of one layer are absent. A neuron can be in three States: waiting, arousal, and refractoriness. Initially, all neurons are in the standby state. When the total charge at the input of the neuron exceeds a certain threshold, it goes into a state of excitation. In this case, a single pulse is generated at the output of the neuron, after which it passes into a state of refractoriness and remains in it for a specified time.

Layers are divided into logical fields of the same size and contain $C = M \times L \times D \times Q$ neurons, where D and Q are the size of the logical field horizontally and vertically, M and L are the number of fields in the layer horizontally and vertically. Information signals in this type of network are distributed in the form of single pulses of excited neurons. The control unit implements spatial shifts of these signals during transmission from one layer to another. The shifts are made by values multiple of D and Q . This is necessary to avoid conflicts between neurons in the network. In addition, due to shifts, the network can be endowed with various logical structures (linear, spiral, loop, etc.).

The information to be processed by the RNN is encoded as a sequence of binary matrices of size $D \times Q$, where the value "0" corresponds to the neuron in the waiting state, and the value "1" corresponds to the excited neuron. Thus, the dimension of the RNN fields is chosen taking into account the possibility of encoding the processed information in the form of a sequence of binary matrices of the appropriate size.

The encoded data is then sequentially "written" to the RNN field, called the "input" field. Submission of the next binary matrix is possible after the neurons of this field will come out of the state of refractoriness, in which they entered after the "record" of the previous matrix on them. In this case, the control unit carries out spatial shifts of information signals in the process of their passage through the RNN, as a result of which they enter the field, called the "output", where they are "read". Due to the presence of an internal logical structure in the RNN, a one-to-one correspondence is established between the input and output signals. In the process of signal propagation within the RNN, the control unit can regulate the degree of associative interaction between the processed signals, neuronal excitation thresholds, the rate of learning weights, as well as accelerate or slow down the internal time of the RNN. Networks of this type are focused on continuous processing of signal sequences (for example, time series, the values of which are encoded in a format perceived by the network).

There are known methods of using RNN with controlled elements for forecasting [20]. This article proposes to evaluate the effectiveness of forecasting demand for goods on the basis of available data on the background of demand, as well as data on the amount of pre-order of goods for 1, 2 and 3 months ahead.

Section 2 of this article describes the structure of the available data and their features, after which the task of the study is formulated. Section 3 describes the method of forecasting time series on a recurrent neural network with controlled elements, as well as methods of forecasting in relation to available data on demand for goods. Section 4 is devoted to the analysis of the results and their discussion. In conclusion, the main points of the study are briefly described, the shortcomings and prospects of the proposed approaches are presented.

2. Structure of initial data and problem statement

The source data is a set of files in Microsoft Excel format. Each file corresponds to a specific product. The demand information is presented in the form of tables, the rows of which correspond to the months of observation. Depending on the product, the number of lines varies from 15 to 48. For example, if there are 24 lines in the file, it means that for the corresponding product the observation was carried out during the last two years. The columns of this table correspond to the numerical index of the month, the volume of real demand for the month, as well as the value of pre-order for 1, 2 and 3 months ahead. For example, if you are considering a 20th month, you will see the volume of real demand for the 20th month, as well as the volume of pre-order for the 21st, 22nd, and 23rd months. Demand values are normalized and lie in the range from 0 to 1. It should be noted that real demand and/or pre-order data may not be available for some months. In such situations, the value of demand in the table is assumed to be 0. In addition, there may be occasional outliers, such as ultra-low or near-peak demand values in individual months, which differ significantly from the overall trend in demand for a given commodity during the observation period.

Based on the available data on the volume of real demand and pre-orders, the task is to forecast real demand for a horizon equal to 4 months.

3. Ways of neural network associative forecasting

Forecasting of real demand with the help of RNN with controlled elements is based on its ability to call information associatively. To do this, its layers are configured by the control unit into a spiral structure. At the level of neural network channels such RNN can be represented as Fig. 1. Promotion of signals from input to output in it is provided by spatial shifts in their transmission from one layer of RNN to another [19].

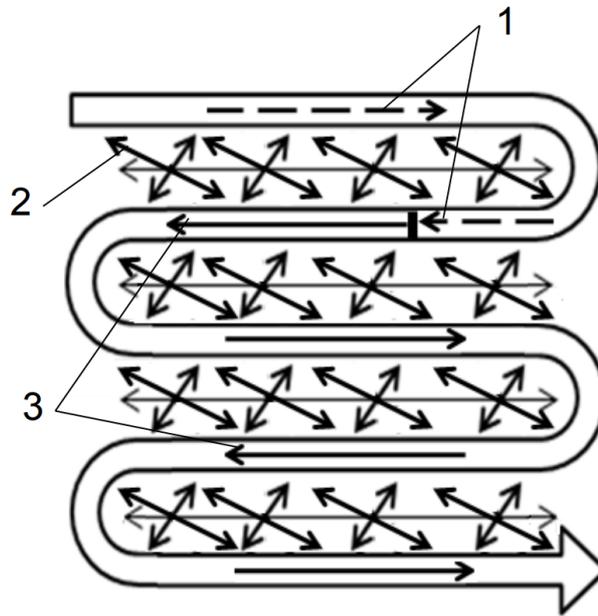


Fig. 1. RNN forecasting scheme with controlled elements. Here 1 is the data called from associative memory (a predicted result), 2 is direction associative call signals from the memory, and 3 is the movement of the original signals.

Available data on demand for goods are encoded in a format perceived by the network, and then fed to its layers. In the process of their passage through the RNN, an associative model is built on its synapses. Before the start of forecasting, the current data supply to the network is interrupted, and the process of associative processing of already received signals is dramatically accelerated. Also, the associative call of signals from memory in the direction of the input is amplified. As a result, the signals processed in the RNN are called future events from memory.

The above method of forecasting can be applied to the existing data structure in two ways. The first way involves encoding all available information into a single sequence and obtaining its predicted values for the 4-month horizon (F_1, F_2, F_3, F_4), as shown in Fig. 2.

In the second way, the prediction occurs in 4 consecutive steps (see Fig. 3). At the first step, the RNN provides data including information about the real demand and pre-order of goods for 1 month in advance. As a result, the time series $\{\text{Real}\}$ is supplemented with the forecast value for the first month (F_1) at the network output. In the second step, this augmented time series is fed back into the neural network, but this time it is combined with pre-order data for the second month. The result is a forecast for the second month (F_2). In the third and fourth steps, the values of F_3 and F_4 are obtained in the same way, with the peculiarity that in the fourth step, the time series " $\{\text{Real}\} F_1 F_2 F_3$ " is not supplemented with pre-order information due to its absence.

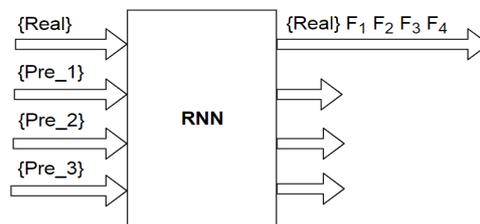


Fig. 2. The first way of forecasting.

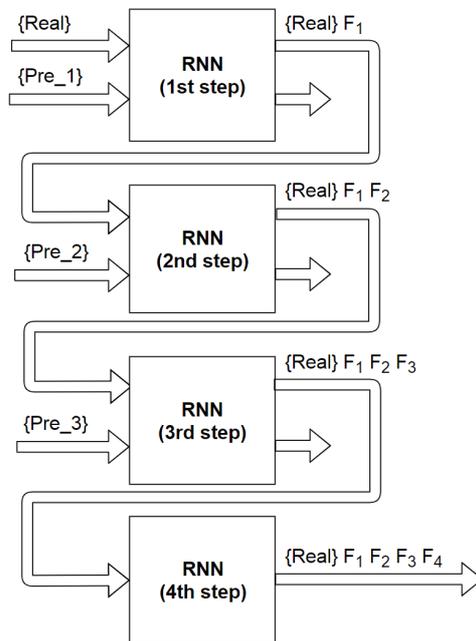


Fig. 3. The second way of forecasting.

The advantage of the first way of forecasting is the ease of implementation and a small number of operations. At the same time, the second way allows to fine-tune the associative model for forecasting.

4. Forecasting results and analysis

To meet the forecasts was configured of a neural network with spiral structure of size 3×5 logical fields. The logical fields were $10 \times 12 = 120$ neurons in size, where the first 100 neurons were assigned to encode the real demand value, the next 20 to encode the pre-order values for the 1st, 2nd and 3rd months. Coding was carried out by the interval method, when each neuron is put in accordance with a certain interval of the total range of possible values of this value. Neuron index is defined as $I = \text{CurrentValue} \times (\text{MaxValue} - \text{MinValue}) / \text{CountOfNeurons}$, where *Current_Value* is the encoded value, *Max_Value* and *Min_Value* are respectively, the maximum and minimum possible values of this value, *Count_Of_Neurons* is the number of neurons allocated for encoding.

Checking the effectiveness of the above ways was performed for 15 items of goods ($i = 1 \dots 15$). In each case, the last 4 values ($k = 1 \dots 4$) were selected for prediction, while the previous ones were used for training the RNN and building an associative model. The obtained values of F_k forecasts were compared with the real values of R_k by the mean absolute percentage error (MAPE), mean absolute error (MAE) and root-mean-square error (RMSE) metrics.

$$MAPE_k = \frac{1}{N} \sum_{i=1}^N \frac{|R_{ki} - F_{ki}|}{R_{ki}} \times 100\%$$

$$MAE_k = \frac{1}{N} \sum_{i=1}^N |R_{ki} - F_{ki}|$$

$$RMSE_k = \sqrt{\frac{1}{N} \sum_{i=1}^N [R_{ki} - F_{ki}]^2}$$

The results of calculations are shown in tables 1 and 2.

Table 1. Prediction accuracy for the first way.

	MAPE	MAE	RMSE
F₁	41.1462	0.1176	0.1597
F₂	56.9379	0.1264	0.1609
F₃	44.4890	0.1700	0.2528
F₄	26.6199	0.0776	0.1231
Average	43.0482	0.1229	0.1741

Table 2. Prediction accuracy for the second way.

	MAPE	MAE	RMSE
F₁	37.5744	0.1062	0.1481
F₂	24.5587	0.0629	0.0924
F₃	31.7938	0.1175	0.1563
F₄	38.4544	0.1110	0.1619
Average	33.0953	0.0994	0.1397

The results of calculations show some advantage of the second way in comparison with the first. Thus, according to the MAPE metric, the accuracy was on average higher on 9.95%, the increase of MAE was 0.0235, and RMSE is 0.0344. It is also worth noting that neither in the first nor in the second case there is a monotonous decrease in the average accuracy of the forecast from F₁ to F₄. Thus, in the first way of forecasting, the maximum average accuracy was achieved for F₄, and in the second way the best result is given for F₂. This phenomenon is explained by the fact that internal regularities in the predicted time series, which due to the associativity of the RNN can be detected by it, do not always correlate with the monthly interval of recording the volumes of real demand. Moreover, depending on the structure of the data and the way of training, different patterns begin to play the greatest role in forecasting. This explains the fact that in the first and second ways of forecasting, the best accuracy of the forecast is obtained for different months.

5. Conclusion

The results of the study on neural network forecasting of demand for goods showed the following. Traditional neural network solutions do not allow achieving high accuracy of forecasts due to their inherent limitations and disadvantages. A new method is proposed based on the use of RNN with controlled elements. In accordance with this

method, two ways of forecasting are proposed. According to the first way, the controlled elements of RNN are trained on all available data and generate forecast demand values for the 4-month horizon. In accordance with the second way, forecasting is carried out in four consecutive steps. At each step, the demand is predicted for the next month, and the training set is supplemented with the values predicted for the previous months. The proposed ways are tested for 15 items of goods, the results of which are compiled tables of efficiency. It is shown that the second way gives a relatively high accuracy, although it requires more operations. The proposed ways can be applied in practice in forecasting demand for various goods.

References

1. Lapygin, Yu.: Economic forecasting. Exmo, Moscow (2009).
2. Zliobaite, I., Bakker, J., Pechenizkiy, M.: Beating the baseline prediction in food sales: How intelligent an intelligent predictor is? *Expert Systems with Applications* 39(1), 806-815 (2012). <https://doi.org/10.1016/j.eswa.2011.07.078>
3. Tsymbalov, E.: Churn Prediction for Game Industry Based on Cohort Classification Ensemble. In: *EEML@CLA*, pp. 94-100. Moscow (2016)
4. Yu, H., Rao, N., Dhillon, I.: Temporal Regularized Matrix Factorization for High-dimensional Time Series Prediction. In: *NIPS*, pp. 847-855. Barcelona (2016)
5. Pham, D., Xing, L.: *Neural Networks for Identification, Prediction and Control*. Springer-Verlag, London (1995). DOI: 10.007/978-1-447 1-3244-8
6. Montavon, G., Orr, G., Müller, K. (Eds.) *Neural Networks: Tricks of the Trade*. 2nd edn. Springer-Verlag, Berlin (2012). DOI 10.1007/978-3-642-35289-8
7. Zhang, X., Hu, L., Zhang L.: An efficient multiple kernel computation method for regression analysis of economic data. *Neurocomputing* 118, 58-64 (2013). <https://doi.org/10.1016/j.neucom.2013.02.013>
8. Kohzadi, N., Boyd, M., Kermanshahi, B., Kaastra, I.: A comparison of artificial neural network and time series models for forecasting commodity prices. *Neurocomputing* 10(2), 169-181 (1996). [https://doi.org/10.1016/0925-2312\(95\)00020-8](https://doi.org/10.1016/0925-2312(95)00020-8)
9. Galeshchuk, S.: Neural networks performance in exchange rate prediction. *Neurocomputing* 172, 446-452 (2016). <https://doi.org/10.1016/j.neucom.2015.03.100>
10. Hu, H., Tang, L., Zhang, S., Wang, H.: Predicting the direction of stock markets using optimized neural networks with Google Trends. *Neurocomputing* 285, 188-195 (2018). <https://doi.org/10.1016/j.neucom.2018.01.038>
11. Srinivasan, D.: Energy demand prediction using GMDH networks. *Neurocomputing* 72 (1-3), 625-629 (2008). <https://doi.org/10.1016/j.neucom.2008.08.006>
12. Wu, W., Wang, X.: The Coal Demand Prediction Based on the Grey Neural Network Model. In: *LISS 2014*. Springer, Berlin, Heidelberg (2015). https://doi.org/10.1007/978-3-662-43871-8_194
13. Cubero, R.: Neural networks for water demand time series forecasting. In: *Artificial Neural Networks. IWANN 1991. Lecture Notes in Computer Science*, vol. 540. Springer, Berlin, Heidelberg (2005). <https://doi.org/10.1007/BFb0035927>
14. Chramcov, B., Varacha, P.: Usage of the Evolutionary Designed Neural Network for Heat Demand Forecast. In: *Nostradamus: Modern Methods of Prediction, Modeling and Analy-*

- sis of Nonlinear Systems. *Advances in Intelligent Systems and Computing*, vol. 192. Springer, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-642-33227-2_13
- 15.Lo, C.: Back Propagation Neural Network on the Forecasting System of Sea Food Material Demand. In: *Advances in Computer Science and Education Applications. Communications in Computer and Information Science*, vol. 202. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22456-0_22
- 16.Herrera-Granda, I., et al.: Artificial Neural Networks for Bottled Water Demand Forecasting: A Small Business Case Study. In: *Advances in Computational Intelligence. IWANN 2019. Lecture Notes in Computer Science*, vol. 11507. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-20518-8_31
- 17.Chawla, A., Singh, A., Lamba, A., Gangwani, N., Soni, U.: Demand Forecasting Using Artificial Neural Networks—A Case Study of American Retail Corporation. In: *Applications of Artificial Intelligence Techniques in Engineering. Advances in Intelligent Systems and Computing*, vol. 697. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-1822-1_8
- 18.Christopher, J., Mou, J., Yin, D.: Convolutional Neural Network Deep-Learning Models for Prediction of Shared Bicycle Demand. In: *International Conference on Applications and Techniques in Cyber Security and Intelligence ATCI 2018. Advances in Intelligent Systems and Computing*, vol. 842. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-98776-7_1
- 19.Osipov, V., Osipova, M.: Space–time signal binding in recurrent neural networks with controlled elements. *Neurocomputing* 308, 194–204 (2018). <https://doi.org/10.1016/j.neucom.2018.05.009>
- 20.Osipov, V.: Neural network with past, present and future time. *Information and control systems* 4, 30–33 (2011).