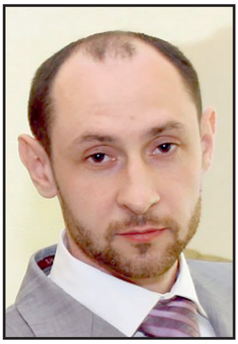
https://doi.org/10.30932/1992-3252-2019-17-3-6-15

Neural Network for Forecasting Energy Consumption Load of a Railway Marshalling Yard






Vladimir N. GRIDIN **Viktor V. DOENIN** **Vladimir S. PANISHCHEV** **Igor D. BYSOV**

Gridin, Vladimir N., *Center for Information Technologies in Design of Russian Academy of Sciences, Moscow, Russia.*
Doenin, Viktor V., *Center for Information Technologies in Design of Russian Academy of Sciences, Moscow, Russia.*
Panishchev, Vladimir S., *Center for Information Technologies in Design of Russian Academy of Sciences, Moscow, Russia.*
Bysov, Igor D., *Southwest State University, Kursk, Russia.**

ABSTRACT

A multilayer neural network has been designed to forecast average daily energy consumption of a railway marshalling yard. The suggested model comprises a multilayer perceptron using 22 inputs, the n-th number of hidden layers and one output. The number of hidden layers in the neural network and neurons in them was chosen experimentally. A comparative selection of activation functions and training methods has allowed for all other parameters to achieve a minimum average relative error.

Two types of loads corresponding to holidays (non-working) and working days were identified. An additional input node with binary coding and two nodes for coding the season were introduced due to a certain repeatability characterizing samples of prediction of loads of energy consumption of the marshalling yard depending on type of a day and on a season. As accounting of the dependence of the forecast on load values in previous days and years (dynamic dependencies) is most important factor, this neural network takes into account the average daily energy consumption during four days of the current period, preceding the forecasted date, and the average daily power consumption during four days prior to this date during last three years.

As a result, considering all factors and experimentally selected parameters of the neural network, the minimum resulting error of MAPE is about 1,4 %, which shows the advantage of the developed neural network in comparison with two other methods of solution of the problem, suggested by other researchers.

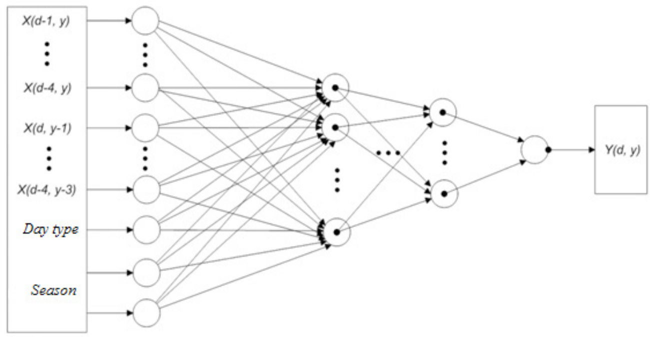
Keywords: multilayer neural network, forecast, perceptron, electric power, load prediction, marshalling yard, railways.

*Information about the authors:
Gridin, Vladimir N. – D.Sc. (Eng), professor, scientific director of the Center for Information Technologies in Design (CITD) of Russian Academy of Sciences, Moscow, Russia, info@ditc.ras.ru.
Doenin, Viktor V. – D.Sc. (Eng), senior researcher of the Center for Information Technologies in Design (CITD) of Russian Academy of Sciences, professor of Russian University of Transport, Moscow, Russia, info@ditc.ras.ru.
Panishchev, Vladimir S. – Ph.D. (Eng), senior researcher of the Center for Information Technologies in Design (CITD) of Russian Academy of Sciences, Moscow, Russia, gskunk@yandex.ru.
Bysov, Igor D. – Master's student of Southwestern State University, Kursk, Russia, bysov93@mail.ru.

Article received 27.02.2019, accepted 30.05.2019.

For the original Russian text please see p. 6.

This work was supported by the Russian Foundation for Basic Research, project 17-20-01133 ofi_m_RZD.



Pic. 1. The structure of the perceptron neural network to forecast energy consumption.

Introduction. Nowadays efficient energy production and power consumption forecasting are among most important topics. The ultimate goal of energy saving is to reduce energy resources costs for an organisation, which necessitates the use of research tools for current consumption monitoring and its forecasting. Regarding railway marshalling yard, this information is one of the elements of the intelligent subsystem for analyzing and forecasting workload.

It is sufficient to use polynomial decompositions [1, p. 42] as a forecasting model for small time intervals, but over a long-time interval the presented models are unsuitable due to impossibility of decomposing data by a basis of small dimension. To achieve sufficient forecast accuracy, autoregressive methods are applicable in that case [2, 3].

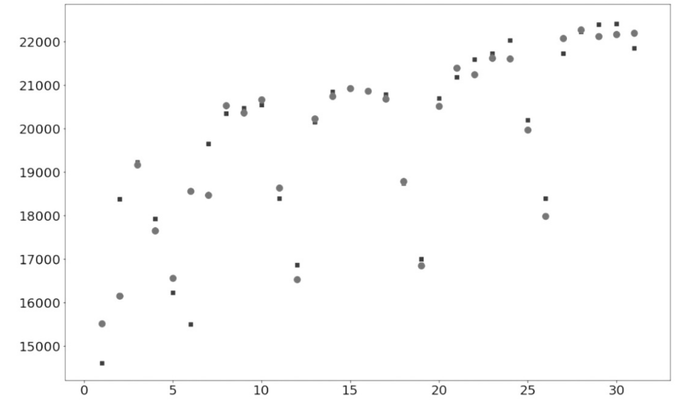
But all the methods listed above have one significant drawback: the application thereof requires direct involvement of an expert. Therefore, it is necessary to apply a more advantageous algorithm that minimizes the human factor when building a forecast. Artificial neural networks that are successfully used to solve forecasting problems can provide such an algorithm [4, p. 17].

The *objective* of the authors is to develop a model of artificial neural network based on a multilayer perceptron to forecast the energy consumption of a railway marshalling yard.

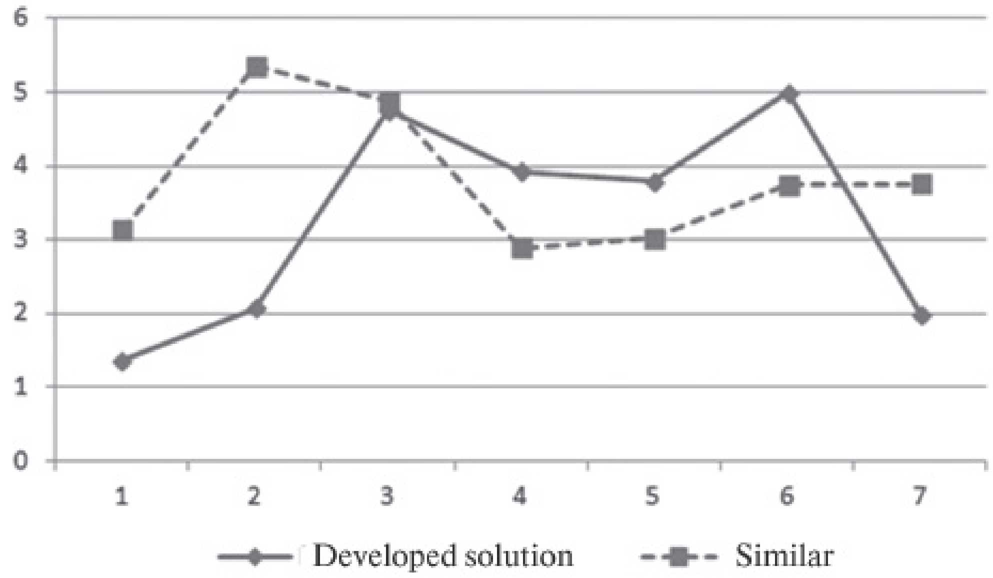
The authors used mathematics and informatics *methods,* programming, particular methods of building artificial neural networks, multilayer perceptron with various activation functions (sigmoid function, rectified linear unit, tanh function), learning methods (L-BFGS, ADAM stochastic optimization algorithm [5]).

Neural network model

To forecast energy consumption load, an array of energy consumption data for 14 years was selected, from which the data



Pic. 2. Graphic image of the neural network.
Rectified linear activation function, L-BFGS training method.



Pic. 3. Comparative graph of forecasted results for the week with the help of the developed neural network and the model described in [10, p. 114].

Table 1

Comparative MAPE error data for weekly forecast

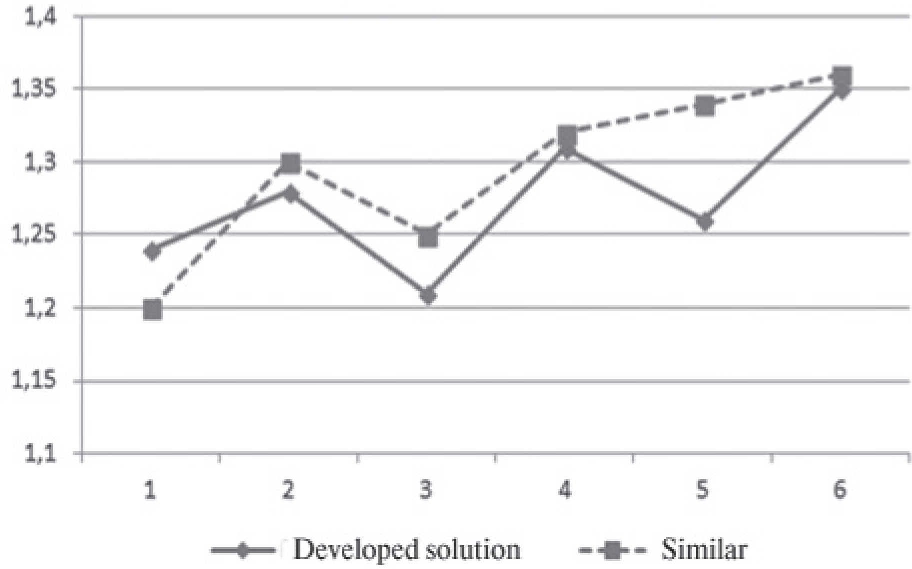
Week-long forecast period	MAPE error in forecasting average daily load with the developed neural network (%)	MAPE error of the model, described in [10, p. 114] (%)
Monday	1.36	3.13
Tuesday	2.07	5.36
Wednesday	4.75	4.87
Thursday	3.92	2.89
Friday	3.79	3.02
Saturday	5.00	3.74
Sunday	1.99	3.76

Table 2

Comparative MAPE error data for a 6-year forecast period

6-year forecast period	MAPE error of the average daily load forecasting with the developed neural network (%)	MAPE error of the model, described in [11, p. 128] (%)
1	1.24	1.2
2	1.28	1.3
3	1.21	1.25
4	1.31	1.32
5	1.26	1.34
6	1.35	1.36

structure of six columns is formed, where column 0 is average daily consumption, 1 – year, 2 – number of the day of the year, 3 – non-working day (0) or working day (1), 4 and 5 – binary coding of the season – winter (11), spring (01), summer (00), autumn (10). Further, an array of the training sample for 12 years is formed as a function dependence of 4 previous days of the current period and 4 relevant previous days of 3 previous years on the type of the day and on the season; besides, leap years are taken into account. Thus, the array has 23 columns, where column 0 contains a dependent variable Y, and 22 columns contain independent variables X. The final steps are purposed to build a neural network in accordance with possible selection of parameters outlined in [6], to train it and to forecast energy consumption for two forthcoming years. The weekends taken into account in this model of the neural network were considered twice as holidays of the Russian Federation and as dates celebrated in Poland for comparison with the analogue that forecasts the loads of the Poland electricity system: January 1, January 6, May 1, May 3, August 15, November 1, November 11, December 25–26, and this must be taken into account to achieve maximum accuracy of forecasting. Most



Pic. 4. Comparative graph of forecasted results for 6 years with the help of the developed neural network and the model described in [11, p. 128].

comprehensive study was carried out using the data of 2002–2015, developing sample of 2005–2013 for training a neural network and a test forecast for the years 2014–2015, that was subsequently verified through comparing its results with actual data. The neural network was built using Python [7–9]. Pic. 1 shows the structure of the developed neural network model.

Selection of neural network parameters

In order to achieve minimum possible forecast error in the developed model of multilayer perceptron, a selection process of optimal activation functions and training methods was carried out. The following neural network parameters were considered:

1. Sigmoid activation function takes on an arbitrary real number at the input, and at the output it gives a real number within the interval from 0 to 1, and so is expressed by the formula: $\sigma(x) = 1/(1+e^{-x})$. As a result of the sigmoid operation, large positive numbers turn into unity, and large negative numbers (modulo) turn into zero.

2. Rectified linear unit activation function (ReLU) is expressed by the formula: $\sigma(x) = \max(0, x)$. Its derivative takes only two values – 0 and 1, which eliminates attenuation or growth of gradients, in addition, the use of a rectified linear function leads to weights decimation, however, it is strongly needed to choose the correct learning speed.

3. Hyperbolic tangent (tanh) activation function takes on an arbitrary real number at the input, and at the output gives a real number in the range from -1 to 1, it can be saturated like the sigmoid, but the output of hyperbolic tangent is centered relative to 0.

4. L-BFGS training (learning) method is a modification of Broyden–Fletcher–Goldfarb–Shanno algorithm, designed to solve nonlinear problems with a large number of unknowns and based on the sequential construction and refinement of a quadratic function model. L-BFGS memorizes the last values of the gradient/function and uses them to make a step using the Newton method and to build a positive definite approximation of the Hessian.

5. Adaptive moment estimation (Adam) training method is a stochastic optimization algorithm that combines the method of a weaker updating of weights for typical features and the method of accumulation of motion. Adam is simple to be implemented, and is computationally efficient algorithm, that has few memory requirements, is invariant to diagonal scaling of the gradients, and is well suited for tasks that are large in terms of data and/or parameters.

Using the above-described training methods and activation functions, an experimental selection of optimal parameters for the developed neural network model was conducted:

1. Rectified linear activation function, L-BFGS training method (Pic. 2). When using one hidden layer with 35 neurons, the average relative error is 1,8 %, the maximum relative error is 29,4 %. When using five hidden layers in the neural network, each of them having from 20 to 30 neurons, the average relative error decreases to 1,4 %, and the maximum relative error to 28,2 %. A further increase in the number of hidden layers and neurons in them does not bring an increase in quality of prediction. When using the adaptive moment estimation training method, the average relative error increases to 2,7 %.

2. Hyperbolic tangent activation function, adaptive moment estimation training method. Single hidden layer with 35 neurons was used, data were normalized. The average relative error was 3,7 %, the maximum relative error was 39,8 %. As the number of layers increases, the error increases.

Using the sigmoid activation function with adaptive moment estimation training method gives an increase in the average relative error to 4 % and an increase in maximum relative error to 41 %.

Tables 1 and 2, Pics. 3 and 4 present a comparative analysis of MAPE errors for the developed perceptron neural network and similar networks.

From Table 1 it can be seen that for a single week, the developed neural network predicts the average daily energy consumption generally better than a similar model.

From Table 2 it can be seen that the highest MAPE error for the developed neural network within 6-year period did not exceed 1,4 %.

Conclusion. A neural network model was developed on the basis of a multilayer perceptron, and an experimental selection of neural network parameters was carried out, which made it possible to predict energy consumption with an average relative error of 1,4 %. Thus, the experimental data confirmed the possibility of using this neural network to forecast energy consumption of a marshalling yard, but it is necessary to take into account the floating holiday schedule to compensate for high prediction error values.

In the course of further development based on the results obtained, this model can also be used to predict energy consumption for a railway station, transport and logistics hub.

REFERENCES

1. Akhmetyanov, R. R., Delegodina, L. A., Kopylova, N. P. [*et al*]. The tasks of energy consumption forecasting in the integrated AMR system of Novosibirsk Scientific Center [*Zadachi prognozirovaniya energopotrebleniya v integrirovannoi ASKUE Novosibirskogo nauchnogo tsentra*]. *Energosnabzhenie*, 2007, Iss. 1, pp. 42–47.

2. Akhmetyanov, R. R., Delegodina, L. A., Kopylova, N. P. [*et al*]. The multiplicative model of the seasonal energy consumption of enterprises [*Multiplikativnaya model sezonnogo energopotrebleniya predpriyatii*]. *Avtometriya*, 2008, Iss. 3, pp. 106–118.

3. Akhmetyanov, R. R., Delegodina, L. A., Kopylova, N. P. [*et al*]. Using non-standard seasonal models of autoregression and integrated moving average in resource saving problems [*Ispolzovanie nestandartnykh sezonnykh modelei avtoregresii i prointegrirovannogo skolzyashchego srednego v zadachakh resursosberezheniya*]. *Avtometriya*, 2008, Iss. 4, pp. 28–41.

4. Bysov, I. D., Panishchev, V. S. Using a neural network to predict the loads of energy consumption of an enterprise [*Ispolzovanie neironnoi seti dlya prognozirovaniya nagruzok energeticheskogo potrebleniya predpriyatiya*]. Tendetsii *razvitiya nauki i obrazovaniya*, 2017, Iss. 33–1, pp. 17–18.

5. Kingma, D. P., Ba, J. Adam: A Method for Stochastic Optimization. Cornell University Library, 2014. [Electronic resource]: https://arxiv.org/pdf/1412.6980.pdf. Last accessed 27.02.2019.

6. Gridin, V. N., Solodovnikov, V. I., Karnakov, V. V. Selection of initial values and optimization of neural network parameters [*Vybor nachalnykh znachenii i optimizatsii parametrov neironnoi seti*]. *Novie informatsionnie tekhnologii v avtomatizirovannykh sistemakh*, 2016, Iss. 19, pp. 270–273.

7. Müller, A., Guido, S. Introduction to machine learning using Python. Trans. from English. Moscow, Williams, 2017, 480 p.

8. Raschka, S. Python Machine Learning. Trans. from English. Moscow, DMK Press, 2017, 418 p.

9. McKinney, W. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython [Russian title: Python and data analysis]. Trans. from English. Moscow, DMK Press, 2015, 482 p.

10. Solomakho, K. L. Application of the principal component method for predicting the volume of energy consumption of a power sales company. Ph.D. (Eng) thesis [*Primenenie metoda glavnykh component dlya prognozirovaniya ob'emov energopotrebleniya energosbytovogo predpriyatiya. Dis... kand. tekh. nauk*]. Southern Urals State University, Chelyabinsk, 2015 [Electronic resource]: https://www.susu.ru/sites/default/files/dissertation/03_text_diss_5.pdf. Last accessed 27.02.2019.

11. Osovsky, S. Neural networks for information processing [*Neironnie seti dlya obrabotki informatsii*]. Moscow, Finansy i statistika, 2002, 344 p.

12. Gridin, V. N., Doenin, V. V., Solodovnikov, V. I., Panishchev, V. S., Trufanov, M. I. On the issue of building an intellectual subsystem for analyzing and predicting the operation of the marshalling yard [*K voprosu postroeniya intellektualnoi podsistemy analiza i prognozarovaniya raboty sortirovochnogo uzla*]. *Informatsionnie tekhnologii i vychislitelnie sistemy*, 2017, Iss. 4, pp. 95–103.

13. Gridin, V. N., Doenin, V. V., Panishchev, V. S. On construction of an intelligent subsystem for analyzing the parameters of a marshalling hub. *World of Transport and Transportation*, Vol. 15, 2017, Iss. 4, pp. 6–19.