УДК 004.832;338.27

O.V. Skakalina, Cand. Sci. (Tech.)

Poltava National Technical Yuri Kondratyuk University, Poltava, Ukraine, e-mail: elena.skakalina@bk.ru

APPLIED ASPECTS OF UTILIZATION OF THE GROUP METHOD OF DATA HANDLING FOR SHORT-TERM FORECASTING

О.В. Скакаліна, канд. техн. наук

Полтавський національний технічний університет ім. Ю. Кондратюка, м. Полтава, Україна, e-mail: elena.skakalina@bk.ru

ПРИКЛАДНІ АСПЕКТИ ВИКОРИСТАННЯ МЕТОДУ ГРУПОВОГО УРАХУВАННЯ АРГУМЕНТІВ ПРИ КОРОТКОСТРОКОВОМУ ПРОГНОЗУВАННІ

Purpose. The creation of the information apparatus for short-term forecasting of key indicators of economic entities (EE) industrial activity based on combinatorial and neural network algorithms of the group method of data handling (GMDH).

Methodology. This work presents a methodology for short-term forecasting of main economic indicators of industrial activity of an economic entity using modifications in the classical algorithm of the GMDH, comparing the effectiveness of combinatorial and neural network algorithms of the GMDH.

Findings. The analysis of forecasting methods reasoned the choice of the neural network algorithm of the GMDH for the research. The optimal model structure has been compiled and the dependence of the output parameters on the selected input parameters of the system has been derived. The analysis of economic indicators of an agricultural holding has been done. The research results are presented via graphic visualization. The effectiveness of the application of the modified GMDH algorithms for forecasting production trends for up to 3 years has been proved. The regularities of the data changes were revealed. The accuracy of the results was evaluated.

Originality. On the basis of the content analysis of the main indicators characterizing the agricultural holding industrial activity, new approaches to the construction of optimal models were proposed, taking into account characteristics of the subject area, which improves the quality of managerial decision making. Information technology for short-term forecasting based on neural network algorithm of the GMDH has been developed.

Practical value. The proposed technique allowed us to conduct the short-term forecasting of main economic indicators of an economic entity industrial activity (selling goods and services, cost of the sold products, gross profit, net profit; commercial, general and administrative expenses). The obtained accuracy assessment of the forecasting results is within the range from 1 to 5%.

Keywords*: information technology, agricultural holding, group method of data handling, forecasting, neural network algorithm, decision support system, economic indicators*

Problem formulation. One of the important indicators of the state economic condition is the gross domestic product (GDP). The problem of identifying factors that ensure formation of stable positive dynamics of real GDP is very important for the world. Therefore, in modern conditions of rapidly changing macroeconomic component, creating new economic conditions, it is necessary to improve methods and systems to control transition to the innovative components of social production and a significant increase in efficiency. These processes provide a detailed scientific justification and optimal decisions that are made in the process of business units' management. Such decisions are grounded on prospective planning and scientific forecasting. High-quality decision-making in economy and finance, as well as in the management of technical systems and technological processes, require high-quality forecasts.

Despite the presence of many forecasting methods, the problem of improving the quality of forecasts is relevant, because the quality depends on many factors, including the quality and amount of experimental data, presence of uncertainty of various kinds and types of the studied processes (stationary and transient, linear and nonlinear). It is very difficult and sometimes impossible to take into account all peculiarities of the processes for a number of reasons: small sample size, big measurements noises, significant perturbations and reductions. It is therefore necessary to search for methods and means to improve the quality of models and forecasts despite these difficulties. To some extent, the task of improving the quality of forecasts is solved by the method of group data handling and neural networks.

Analysis of recent research and publications. The decision support system (DSS) is widely used to improve the efficiency of regional management, via construction, analysis and application of mathematical models of regional socio-economic systems or individual subsystems.

One of the characteristic inherent problems related to mathematical modeling of socio-economic systems is the presence of a large number of factors influencing the study of a socio-economic indicator (output model) and limited (regarding the number of factors) amount of observations of these factors used in the building model. To solve such problems, academician A.G. Ivakhnenko [1] proposed the method of group data handling (GMDH) that ensures acceptable quality of the models in terms of multifactorial controlled object and the limited scope of the training set.

However, in making decisions on management of socio-economic systems, one must take into account a large number of interdependent parameters and control impacts, many of which can not be measured with a metric scale. It should also be noted that in conditions of a weak mathematical formalization of some socio-economic processes and the limited amount of statistics, the role of expert information in decision making increases. These circumstances reduce the efficiency of the known applications of GMDH as a part of mathematical and algorithmic support for DSS management of socio-economic systems at the regional level [2]. To minimize these shortcomings of GMDH, the authors proposed a method based on the use of fuzzy-logic (hybrid) neural networks in the construction of models. In [3], the authors put forward a fuzzy method of group data accounting with fuzzy input. It is a further development of fuzzy method of group data handling. The method allows to build the relationship between fuzzy economic indicators and output economic and ecological indicator. As a result of the method application, confidence interval for the output variable is obtained, which is clearer than obtaining point estimation. The advantage of this method is the fact that we should not worry about the type of predictive model – the algorithm builds a model automatically. In [4] the authors analyzed two basic methodological approaches to prediction. First approach is called genetic (also known as descriptive). According to it, the assessment is made on the basis of the research into the causal relationships between effects. Genetic approach suggests the possibility to formulate any plausible, scientifically grounded quantitative and qualitative, general and partial hypotheses regarding future factors and conditions development, including the hypothesis of economic policy definition and other forms of active conscious impact on production and economic processes. Another approach is called regulatory, or targeted approach. It is based on determining the results to be achieved in the future.

According to the first approach, relationship and sequence of phenomena are treated from the present to the future, in the second – from the future to the present.

The latter perspective analyses the chain of events that may occur and measures to be taken to achieve a given regulatory result in the future.

Regulatory/targeted approach has number of common features with the genetic approach. Developing predictions for an enterprise involves application of different specific methods: method of expert assessment, methods of logical modeling, mathematical modeling, conventional method. Typically, economic forecasting incorporates a combination of methods. Analysis of GMDH methods combinations and artificial neural networks as a method of forecasting is given in [5].

Components of the general problem to be solved. Creation of economic models in order to obtain reliable forecasting data using neural networks as a form of GMDH is the most topical trend of modeling and forecasting, which has a great capacity and opportunities for further development. Application of GMDH for forecasting economic performance is an urgent task, as is evident from by recent publications. However, despite the study, the problem of forecasting has not been solved yet, as the number of factors that influence the behavior of the predicted process usually approaches infinity.

Statement of the problem. The aim of this work is to analyze existing approaches and methods for forecasting major economic indicators of agricultural holding (AH), development of information technology for forecasting revenues and production costs during the next three years using as an example AH, examine patterns of data change, and to assess the accuracy of the results. Method of group data handling (GMDH) is used in various fields for data analysis, forecasting and systems modeling, optimization and pattern recognition. Inductive GMDH algorithms provide a unique opportunity to automatically discover interdependencies in data, choose the optimal model or network structure, improve accuracy of existing algorithms. This approach to models self-organization fundamentally differs from deductive methods that are commonly used. It is based on inductive principles when solution is grounded on iterating of external criterion. With different iterating solutions, inductive modeling methods attempt to minimize the role of the author's assumptions in the modeling results. The computer itself finds a model structure and the laws governing the facility. It can be used in the creation of artificial intelligence as an advisor for dispute resolution and decision making. Method of group data handling consists of several algorithms for solving various problems. It includes both parametric and clustering algorithms, combining analogues, rebinarization and probabilistic algorithms. This approach of self-organization is based on iterationof models that gradually become more complex, and on choosing the best solution according to the minimum of external criterion. Not only polynomials but also non-linear, probabilistic functions or clusters are used as the basic models. The advantages of GMDH comprise:

• It helps to find the optimal complexity of the model structure, adequate to the level of noise in the sample data. (For solving real problems with noisy or short data, simplified forecasting models seem to be more accurate).

• The number of layers and neurons in the hidden layers, structure and other optimal parameters of neural networks are found automatically.

• Finding the most accurate or unbiased model is guaranteed because the method does not miss the best solution in the iteration of all options (in a given class of functions).

• Any non-linear functions or factors that may influence the original variable are used as input arguments.

• The method automatically finds interpretable interrelationships in data and selects effective input variables.

• GMDH algorithms iterations are simple to program.

• GMDH networks are used to improve the accuracy of other algorithms modeling.

• The method uses the information directly from the sample data and minimizes the effect of a priori assumptions about simulation results.

• It enables to find an unbiased physical model of an object (law or cluster) – the same for all future samples.

Rather general formulation of the problem of structural-parametric identification, or building of models from experimental data, can be reduced to the search for some extremum of criterion *CR* in innumerable different models ζ.

$$f^* = arg\min_{f\in\zeta} CR\,(f) \quad . \qquad (1)$$

It is obvious that (1) does not provide comprehensive formulation of the problem, thus it is necessary to: specify the type and volume of initial information; specify the class of basic functions (operators) which formed the set ζ; determine the way to generate models *f* together with parameter estimation method; select a models comparison criterion; specify the method of minimizing *CR*. Let us clarify this statement, assuming that a given sample $W = [X; y]$, contains n points of observations that form the matrix $X = \{x_{ij},\ i = l, ..., n;\ j = l, ..., m\}$ and vector $y = (y_l ... y_m)^T$, when $n \geq m$.

In general, the process of solving the problem of structural-parametric identification (1) includes primarily the following stages:

1. Setting the data sample (obtained as a result of passive or active experiment) and a priori information.
2. Selecting or setting the class of basic functions and data conversion.
3. Generating different models structures in the selected class.
4. Evaluating parameters of the generated structures and forming *F* set.
5. Minimizing the specified criterion *CR (f)* and selecting the optimal model *f.*
6. Checking adequacy of the obtained optimal model.
7. Making a decision about the process completion.

These stages describe the random process of building models. Depending on prior information and the purpose of modeling, certain stages may be omitted. For example, in the case of parametric identification problem, set *F* consists of one model (one set structure), that is the 5-th stage of the process is omitted.

On the whole, the problem of identification boils down to building some set of models of different structure types according to the sample data

$$\hat{y}_f = f(X, \hat{\theta}_f) \qquad (2)$$

and finding the optimal model on condition that

$$f^* = arg\min_{f\epsilon F} CR\left(y, \left(X, \hat{\theta}_f\right)\right), \qquad (3)$$

while parameters assessment for each model $f\epsilon F$ is the solution for another extreme problemof the type

$$\hat{\theta}_f = arg\min_{\theta_f\epsilon R^m} QR(y, X, \theta_f), \qquad (4)$$

where θ_f is called *the complexity* of the model *f* and is equal to the number of non-zero components in the model of the type (2,3) and *QR* − quality criterion for solving the problem of parametric identification of each private model that is generated in the task of structural identification.

Most GMDH algorithms use a polynomial basis function. The general relation between the input and output variables can be expressed as Volterra function series which is a discrete analogue of Kolmogorov-Gabor polynomial

$$y = a_0 + \sum_{i=1}^{M} a_i x_i + \sum_{i=1}^{M}\sum_{j=1}^{M} a_{ij} x_i x_j + \\ + \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M} a_{ijk} x_i x_j x_k\ , \qquad (5)$$

where $x = (x_1, x_2, ..., x_M)$ are input vector variables; $A(a_1, a_2, ..., a_M)$ vector of coefficients or weights.

Independent variables, functional forms or final difference members can be the components of the input vector *X*. Other nonlinear basis functions, such as differential, logistics, probabilistic or harmonic can also be used to build the model. The method allows to simultaneously obtain the optimal model structure and the dependence of the output parameters on the selected most important input parameters of the system [6].

The basic material. Two GMDH algorithms have been used in the research − combinatorial and neural network algorithms. Combinatorial GMDH model is a polynomial function which is linear in terms of parameters. Combinatorial model is a subset of polynomial functions conditions generated from a given set of variables. For example, if the data model is a set of two input variables x_1 and x_2 and output (target) variable y, quadratic polynomial is commonly used, which requires the following optimization

$$y = a_0 + a_1{\cdot}x_1 + a_3{\cdot}x_1{\cdot}x_2 + a_4{\cdot}x_1^2 + a_5{\cdot}x_2^2\,. \qquad (6)$$

The power of polynomial functions is defined by the user, and can be a linear function, that is with power equal to one. Combinatorial GMDH algorithm selects the optimum-complex model, for example: $y = a_0 + a_3 \cdot x_1 \cdot x_2$ as a subset of the full polynomial with the smallest error in the data model. The stage of preliminary data processing allows using different operators of variables $x_1$and x_2, for example, exponential, sigmoid function, time series charts, etc. However, the final model will still be linear in terms of parameters. The complete combinatorial search for the model components often takes too long, so we can limit it by searching models with no more than n members. For example, models with only 2 conditions allow to search among 100, 000 variables and even larger sets. At the same time, the complete search is not recommended for model spaces with more than 25 linear or polynomial conditions. Combinatorial GMDH algorithm is usually time-consuming. Its application as a self-governing algorithm is practically useful for selecting variables or difficulties with the limit of 3–7 of the model parameters that can handle 500–100 variables.

GMDH neural network algorithm. GMDH-like neural networks, also known as polynomial neural networks, use combinatorial algorithm for optimizing neural communication. The algorithm iteratively creates layers of neurons with two or more inputs. The algorithm saves only a limited set of optimal complex neurons which are defined as the initial width of the layer. Each new layer is created using two or more neurons taken from any of the previous layers. Each neuron in the network uses the transfer function (usually with two variables), allowing for the combinatorial search to choose the transfer function which most accurately predicts the data. The transfer function is usually of quadratic or linear form.

GMDH-like network create many layers, but layers of compounds are as rare as few joints in one layer.

As mentioned above, the algorithm returns only a limited number of neurons in each layer; as each new layer can be connected to the previous layer, the width is constantly growing. Given how rarely the upper layers of the population improve the number of models, we divide the additional size of the next layer by two and generate only half the neurons derived from the previous layer, ie the number of neurons N in layer k is$k$ equals $N_k=0.5{\cdot}N_{k-1}$. This makes heuristic algorithm faster, while the chance to diminish the quality of models is rather low.

Table 1

Financial Indices and their Components

The financial result	The financial index
Sale of goods and services	1. Own production 2. Goods for resale 3. Other sales of goods and services 4. Quantity discounts 5. Provision for goods return
Cost of sales	1. Raw materials 2. Wages 3. Depreciation 4. Costs for utilities 5. Cost of goods for resale 6. Expenses on pension benefits 7. Other
Grossprofit	Sale of goods and services + Cost of sales
Selling, general and administrative expenses	1. Wages (excluding pension costs) 2. Transportation 3. Pension costs 4. Taxes (other than income tax) 5. Materials and supplies 6. Protection 7. Advertising and marketing 8. Depreciation of tangible and intangible assets 9. Auditing, consulting and legal services 10. Charging provision for doubtful debts 11. Public services 12. Services in the field of information technology and communication 13. Veterinary services 14. Insurance 15. Banking services 16. Repair and maintenance 17. Other
Other operating (income) expenses	1. Loss on disposal offixedassets 2. Extraordinarylossesassociatedwiththeprivatizationof a subsidiary
Profit from operations	Gross profit + Selling, general and administrative expenses + Other operating (income) expenses
Other operating (income) expenses	1. Percent costs 2. Loan fees 3. The fee for early repaymen of EBRD loan 4. (Profit) / loss on foreign exchange differences 5. Gain on debt relief 6. Gain on early redemption of bonds 7. Other financial (income) and expenses, net
Profit from continuing activities before tax, minority interests and extraordinary income	Income from operating activities + Other income and expenses, net
Income tax	
The share of minority shareholders	
Net profit	Profit from continuing activities before tax, minority interests and extraordinary income + Income taxes + minority interests

Generation of new layers usually stops, when a new layer cannot demonstrate a testing accuracy better than the previous layer. However, creation of new layers also stops if the testing error was reduced by less than 1%, or if the number of layers has reached a specified limit.

The analysis of input data. Financial results of the company performance are characterized by growth in the amount of equity (net assets), which is the main source of profit from operation, investment, fiscal activities, and of the profit resulting from extraordinary circumstances.

Profit is a part of net income that the company obtains directly after sales as a reward for the invested capital and the risk of entrepreneurship. Quantitatively, it is the difference between the total revenues (after payment of value added tax, excise tax and other deductions from revenue to the budget and non-budget funds) and total expenses of the reported period (Table 1). The amount of profit and profitability level depend on production, supply, marketing, rubble, investment and financial activities of the company. Therefore, these indicators describe all aspects of management.

Practical implementation and analysis of the results. Prediction of the main types of profits and expenses. Forecasting time series was performed in the GMDH Shell software environment. GMDH Shell is a simple but powerful software for data mining and forecasting based on GMDH. GMDH Shell allows to explore data, build a regression model, and apply the previously obtained model for forecasting.

Model building by GMDH is carried out in two stages:

Stage 1 – application of GMDH for the data obtained through programs for forecasting of output parameter;

Step 2 – forecasting input parameters using the method of autoregressive integration of the average variable (ARIMA - Autoregressive integrated moving average) [7]. The forecast horizon is assumed to be 3 years.

1. Combinatorial (Fig. 1)
2. GMDH Neural Network (Fig. 2)

As seen from the results (Table 2), the percent of deviations varies from 1 to 5, which, considering such a small sample, can be taken as quite a good yield. The deviations values also show that they are kept within acceptable norms.

Comparing the results of both forecasting algorithms we can conclude that the algorithm of GMDH neural network is more efficient because the percentage deviations are lower than those within the combinatorial method. Therefore, we will use this algorithm in the following forecasts.

Predicting major economic indicators of the economic entity. The first step of forecasting is to choose the data and the variables to be forecasted (Fig. 3, 4).

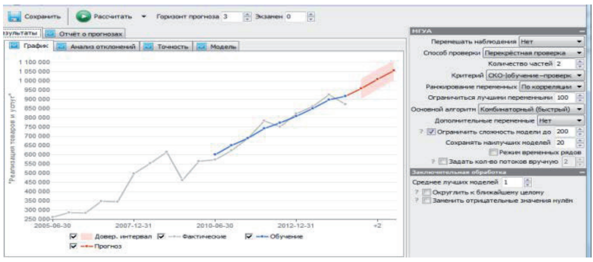

Fig. 1. Prediction made by combinatorial algorithm

Fig. 2. Prediction made by neural network algorithm

Table 2

Comparison of the Prediction Results

PREDICTION RESULTS BY COMBINATORIAL ALGORITHM								
ID	2012-06-30	2012-12-31	2013-06-30	2013-12-31	2014-06-30	+1	+2	+3
ACTUAL	749255,00	821064,00	859560,00	925359,00	872256,00			
FORECAST	770515,14	808739,53	850506,72	896760,96	916470,46	957350,20	1008491,90	1056324,24
DEVIATION	21260,14	-12324,47	-9053,28	-28598,04	44214,46			
% DEVIATION	2,84	-1,50	-1,05	-3,09	5,07			
PREDICTION RESULTS BY NEURAL NETWORK ALGORITHM								
ACTUAL	749255,00	821064,00	859560,00	925359,00	872256,00			
FORECAST	756894,43	818752,25	835603,64	891776,68	915272,08	973408,49	978034,69	1114766,97
DEVIATION	7639,43	-2311,75	-23956,36	-33582,32	43016,08			
% DEVIATION	1,02	-0,28	-2,79	-3,63	4,93			

The results of the forecast of the main economic indicators is presented in the Fig. 5–8.

The forecast horizon is 3 years. The algorithm used – GMDH Neural Network. The obtained results:

- ✓ Cost of sales.
- ✓ Gross profit.
- ✓ Net income.

Thus, we can conclude that by December 31, 2015:

- ✓ Production costs will increase by 13.9%.
- ✓ Gross profit will increase by 16.2%.
- ✓ Net income will increase by 10.2%.
- ✓ Sales of goods and services will grow by 27.8%.

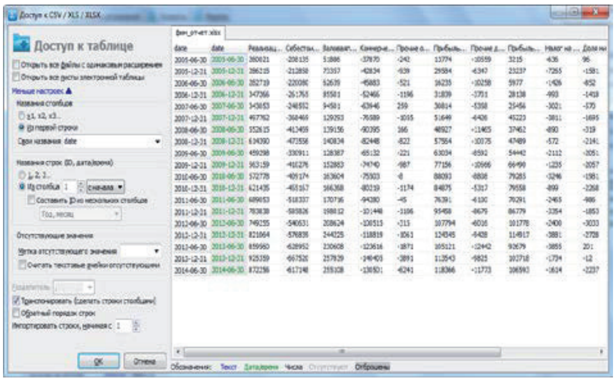

Fig. 3. Selecting data table

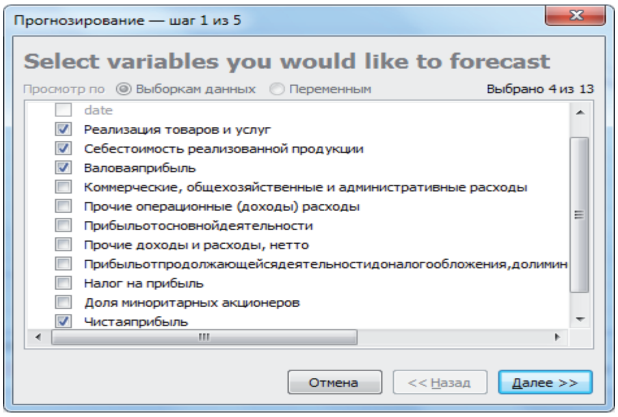

Fig.4 .The choice of variables

Fig. 5. Projected schedule of cost of sales

Prediction of the main types of costs. The main cost of production is defined by commercial, general and administrative expenses (Fig. 8). Table 3 shows that next year expenditures will increase by 16.5%. The projected costs were compared with previous data (Table 4).

Conclusions and suggestions. The effective operation of an enterprise in modern conditions is impossible without the use of management techniques in forecasting. Prediction per se is an effective tool in making reasonable management decisions, providing probabilistic assumptions about possible future state of the enterprise. Availability of the forecast helps to avoid wrong decisions, and prevent unwanted consequences. Thus, forecasting is a means of developing management strategies for AH. The analysis of the main types of income and expenses for AH manufacture has been conducted on 30 June 2014.

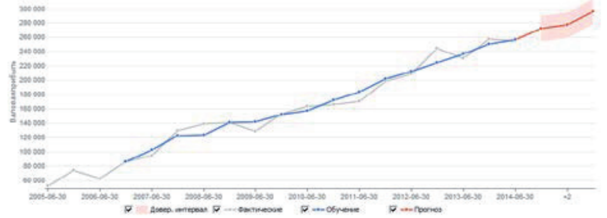

Fig.6. Projected gross profit graph

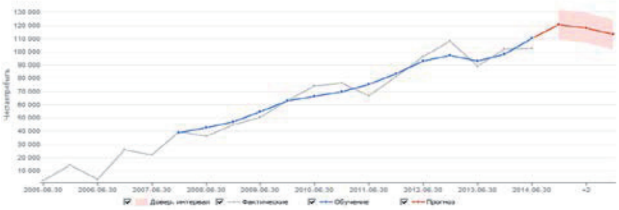

Fig. 7. Projected net income schedule

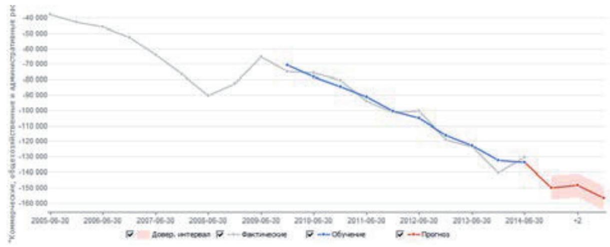

Fig. 8. Graph of forecasting selling, general and administrative expenses

On the basis of the relevant combined and consolidated interim statements of income forecasting, we have conducted a forecasting of major revenue profits and selling, general and administrative costs. The following tasks have been solved:

1. The data for forecasting have been collected and systematized.
2. The collected data have been analyzed.
3. The data have been interpreted as graphs and charts.
4. The procedure of time series prediction has been conducted.
5. The efficiency and accuracy of forecasting has been analyzed.

The developed forecast based on neural network GMDH algorithm proved its efficiency, which indicates reliability of the obtained data. Prospects for further development are related to the improvement of the decisions quality at the DSS by way of transfer to the innovative components of social production and a significant increase in efficiency. These processes ensure a thorough scientific justification, optimal decision-making in management of business entities. Forward planning and scientific forecasting. Forward planning and scientific forecasting form the basis for such decisions.

Table 3

Summary of the Prediction Results

Results of predicting cost of sales								
ID	2012-06-30	2012-12-31	2013-06-30	2013-12-31	2014-06-30	+1	+2	+3
ACTUAL	-540631,00	-576839,00	-628952,00	-667520,00	-617148,00			
FORECAST	-542742,98	-575532,27	-638249,42	-632305,32	-621516,72	-668951,84	-682330,24	-703285,92
DEVIATION	-2111,98	1306,73	-9297,42	35214,68	-4368,72			
%DEVIATION	0,39	-0,23	1,48	-5,08	0,71			
Results of predicting gross profit								
ACTUAL	208624,00	244225,00	230608,00	257839,00	255108,00			
FORECAST	212501,54	235010,99	236704,68	250826,16	256737,85	272138,80	277509,74	296477,19
DEVIATION	3877,54	-9214,01	6096,68	-7012,84	1629,85			
%DEVIATION	1,86	-3,77	2,64	-2,72	0,64			
Results of predicting netprofit								
ACTUAL	96345,00	108308,00	89025,00	101972,00	102742,00			
FORECAST	93168,96	97520,27	93109,21	98208,13	107284,16	120877,56	118231,73	113259,56
DEVIATION	-3176,04	-10787,73	4084,21	-3763,87	4542,16			
%DEVIATION	-3,30	-9,96	4,59	-3,69	4,42			
Results of predicting costs								
ACTUAL	-100515,00	-118819,00	-123616,00	-140405,00	-130501,00			
FORECAST	-105087,58	-115974,93	-122698,11	-132594,44	-133512,68	-149874,74	-148539,30	-151579,41
DEVIATION	-4572,58	2844,07	917,89	7810,56	-3011,68			
%DEVIATION	4,55	-2,39	-0,74	-5,06	2,31			

Table 4

Projected Costs and their Increase Compared with Previous Data

Costs	30.06. 2014	31.12. 2014	30.06. 2015	31.12. 2015	%
Wages (excluding pension costs)	50 213	54492,25	59783,46	62844,82	19,18
Transportation Services	14 819	16608,52	16644,90	18864,45	13,80
Pension expense	9 261	9621,08	10971,03	11687,31	15,40
Taxes (except tax on profits)	6 803	6639,46	7429,98	6434,66	1,94
Materials and supplies	6 677	7160,99	7375,31	7285,96	9,12
Security Guard	5 865	7754,77	8358,61	8592,28	19,22
Advertising and Marketing	5 430	3769,15	3201,92	2165,93	4,34
Amortization of means of production and intangible assets	5 219	4103,11	3452,15	5293,16	5,25
Audit, consulting and legal services	1 958	3152,65	2018,69	3193,33	12,02
Costs (recovery) of doubtful accounts	1 864	2165,73	1104,56	2895,63	39,25
Utilities	1 861	2490,11	2650,48	2920,09	24,67
Services in the Information Technology and Communications	1 558	1455,14	1647,54	2031,77	23,99
Veterinary services	1 528	1777,39	1783,10	1901,12	24,42
Insurance	1 242	1310,58	808,71	924,01	6,60
Banking services	652	1188,03	826,14	1066,41	31,35
Repair and Maintenance	648	1045,84	783,69	998,94	23,29
Others	14 903	16397,06	17370,59	18394,57	23,43
The overall increase in costs					*16,54*

References / Список літератури

1. Ivakhnenko, A.G. and Yurachkovskiy, Yu.P. (1987), *Modelirovaniye slozhnykh sistem po eksperimentalnym dannym* [Modeling of Complex Systems Based on Experimental Data], Radio i Sviaz, Moscow 120 p.

Ивахненко А.Г. Моделирование сложных систем по экспериментальным данным / Ивахненко А.Г., Юрачковский Ю.П. – М.: Радио и связь, 1987. – 120 с.

2. Boyarinov, Yu.G. (2006), "Use of neuro-fuzzy group method of accounting arguments for building models of regional socio-economic systems", *Informatsyonnye Tekhnologii Modelirovaniya i Kontrolya*, Vol. 4(29), pp. 6−12.

Бояринов Ю.Г. Использование нейро-нечеткого метода группового учета аргументов для построения моделей региональных социально-экономических систем / Ю.Г. Бояринов // Информационные технологии моделирования и управления. − 2006. − Вып. 4(29). – С. 6−12.

3. Zaichenko, Yu.P. and Zaiets, I.O. (2007), "Comparative analysis of the predictive models built using fuzzy algorithms and fuzzy GMDH using different algorithms for generating fuzzy prediction models", *Proc. of the Intern. Conf. on Inductive Modelling*, Kyiv: National Academy of Sciences of Ukraine, no.4, pp. 158–165.

Зайченко Ю.П. Порівняльний аналіз прогнозуючих моделей, побудованих за допомогою нечітких алгоритмів МГУА з використанням різних алгоритмів генерації нечітких прогнозуючих моделей / Ю.П. Зайченко, І.О.Заєць // Матеріали міжнар. Семінару з індуктивного моделювання. – К.: НАН України, МННінформ. технол. Та систем, 2007. − №.4. − С. 158−165.

4. Biduk, P.I., Hasanov, A.S. and Vavilov, S.Ye. (2013), Analysis of the Quality Assessments of forecasts using the method of complexation, *Systemni Doslidzhennia ta Informatsiini Technolohii*, no.4, pp. 7−16.

Бидюк П.И. Анализ качества оценок прогнозов с использованием метода комплексирования / П.И. Бидюк, А.С. Гасанов, С.Е. Вавилов // Системні дослідження та інформаційні технології – 2013. − №4. − С. 7−16.

5. Sineglazov, V.M., Chumachenko, Ye.I. and Gorbatyuk, V.S. (2012), "Method of solving the problem of forecasting based on a variety of estimates", *Induktyvne Modeliuvannia Skladnykh System*, Vol. 4, pp. 214−223.

Синеглазов В.М. Метод решения задачи прогнозирования на основе комплексирования оценок / В.М. Синеглазов, Е.И. Чумаченко, В.С. Горбатюк // Індуктивне моделювання складних систем. – 2012. − Вип. 4. − С. 214−223.

6. Litvinov, V.V. and Zadorozhny, A.A. (2012), "Creating block models of systems and processes using the group method of data handling", *Matematychni Mashyny i Systemy*, no.2, pp. 107−116.

Литвинов В.В. Создание блочных моделей систем и процессов с использованием метода группового учета аргументов / В.В. Литвинов, А.А. Задорожний // Математичні машини і системи. – 2012. − №2. − С.107 −116.

7. GMDH Shell, available at: http://www.machinelearning.ru/GMDH_Shell

Мета. Створення інтелектуальної інформаційної технології короткострокового прогнозування основних показників виробничої діяльності господарських суб'єктів (ГС) на основі комбінаторного та нейромережевого алгоритмів методу групового урахування аргументів.

Методика. У роботі представлена інтелектуальна інформаційна технологія короткотермінового прогнозування основних економічних показників виробничої діяльності господарського суб'єкта на прикладі агрохолдингу (АХ) із застосуванням модифікацій класичного алгоритму методу групового урахування аргументів (МГУА), представлене порівняння ефективності комбінаторного та нейромережевого алгоритмів МГУА.

Результати. Зроблено аналіз методів прогнозування та обґрунтовано вибір нейромережевого алгоритму МГУА при виконанні досліджень. Складена оптимальна структура моделі та залежність вихідних параметрів від обраних вхідних параметрів системи. Виконано аналіз економічних показників АХ . Зроблена графічна візуалізація результатів досліджень. Доведена ефективність застосування модифікацій алгоритмів МГУА для прогнозування виробничих трендів на термін до 3-х років. Виявлені закономірності зміни даних. Розрахована оцінка точності отриманих результатів.

Наукова новизна. На основі змістовного аналізу основних показників виробничої діяльності АХ запропоновані нові підходи до побудови їх оптимальних моделей з урахуванням особливостей предметної області, що дозволяє підвищити якість прийняття управлінських рішень, удосконалити алгоритмічне забезпечення систем підтримки прийняття рішень (СППР) щодо керування агрохолдингами. Розроблена інформаційна технологія короткотермінового прогнозування на основі нейромережевого алгоритму МГУА.

Практична значимість. На основі запропонованої технології було проведене короткотермінове прогнозування основних економічних показників виробничої діяльності господарського суб'єкта (реалізація товарів та послуг, собівартість реалізованої продукції, валовий прибуток, чистий прибуток, комерційні, загальногосподарські та адміністративні витрати). Отримана оцінка точності результатів прогнозування, що знаходиться в інтервалі від 1 до 5%.

Ключові слова: *інформаційна технологія, агрохолдинг, метод групового врахування аргументів, прогнозування, нейромережевий алгоритм, системи прийняття рішень, економічні показники*

Цель. Создание интеллектуальной информационной технологии краткосрочного прогнозирования основных показателей производственной деятельности хозяйственных субъектов (ХС) на основе комбинаторного и нейросетевого алгоритмов метода группового учета аргументов.

Методика. В работе представлена интеллектуальная информационная технология краткосрочного прогнозирования основных экономических показателей производственной деятельности хозяйственного субъекта на примере агрохолдингу (АХ) с применением модификации классического алгоритма метода группового учета аргументов (МГУА), представлено сравнение эффективности комбинаторного и нейросетевого алгоритмов МГУА.

Результаты. Сделан анализ методов прогнозирования и обоснован выбор нейросетевого алгоритма МГУА при выполнении исследований. Составлена оптимальная структура модели и зависимость исходных параметров от избранных входных параметров системы. Выполнен анализ экономических показателей АХ. Сделана графическая визуализация результатов исследований. Доказана эффективность применения модификации алгоритмов МГУА для прогнозирования производственных трендов на срок до 3-х лет. Обнаружена закономерность изменения данных. Рассчитана оценка точности полученных результатов.

Научная новизна. На основе содержательного анализа основных показателей производственной деятельности АХ предложены новые подходы к построению их оптимальной модели с учетом особенности предметной области, что позволяет повысить качество принятия управленческих решений, усовершенствовать алгоритмическое обеспечение системы поддержки принятия решений (СППР) относительно управления агрохолдингами. Разработана информационная технология краткосрочного прогнозирования на основе нейросетевого алгоритма МГУА.

Практическая значимость. На основе предложенной технологии было проведено краткосрочное прогнозирование основных экономических показателей производственной деятельности хозяйственного субъекта (реализация товаров и услуги, себестоимость реализованной продукции, валовый доход, чистая прибыль, коммерческие, общехозяйственные и административные расходы). Получена оценка точности результатов прогнозирования, которая находится в интервале от 1 до 5%.

Ключевые слова: *информационная технология, агрохолдинг, метод группового учета аргументов, прогнозирование, нейросетевой алгоритм, система принятия решений, экономические показатели*

Рекомендовано до публікації докт. техн. наук О.Л. Ляховим. Дата надходження рукопису 25.12.14.

Gongwen Xu[1],
Xiaomei Li[2],
Jian Lei[3],
Honglan Zhou[3]

1 – Shandong Jianzhu University, Jinan, Shandong, China
2 – Shandong University, Jinan, Shandong, China
3 – Dongying Branch, Dongying, Shandong, China

AN IMAGE RETRIEVAL AND SEMANTIC MAPPING METHOD BASED ON REGION OF INTEREST

Гунвень Сюй[1],
Сяомей Лі[2],
Цзянь Лей[3],
Хунлань Чжоу[3]

1 – Шаньдунський університет будівництва та архітектури, м. Цзінань, провінція Шаньдун, КНР
2 – Шаньдунський університет, м. Цзінань, провінція Шаньдун, КНР
3 – Дун'їнська філія Шаньдунського інституту технічного контролю спецприладів, м. Дун'їн, провінція Шаньдун, КНР

МЕТОД ПОШУКУ ТА СЕМАНТИЧНОГО ВІДБОРУ ЗОБРАЖЕНЬ ЗА ОБЛАСТЮ ЗОБРАЖЕННЯ, ЩО ПРЕДСТАВЛЯЄ ІНТЕРЕС

Purpose. In order to increase the precision and recall rate of image retrieval method, a new image retrieval method is proposed, which bases on the idea that different parts of an image have different discriminative capabilities and contribute to the image retrieval to different extents. The region of interest method was used to improve the effectiveness of the semantic mapping of content and image retrieval.

Methodology. Based on the region of interest, we got a semantic class range in the feature space by semi-supervised learning. The semantic class range is the mapping of image and class. In view of the different image features, the most suitable distance measurement was chosen.

Findings. The region of interest was divided from image by the improved Harris detection method and extract the low-level image features. The features in database were used as training data for semi-supervised learning to build the mapping of images and semantic classes which is updated iteratively by the relevance feedback.

Originality. The map between an image region of interest and class labels was studied. Two kinds of distances were applied to calculate the different feature's similarity. The semantic mapping is updated in real time. The research realizes a novel and comprehensive image retrieval system.

Practical value. Different needs of users were considered, and the very good experiment results have been received. The image retrieval method proposed in this paper is quicker and more effective than other methods.

Keywords*: non-subsampled contourlet transform, semantic gap, relevance feedback, image retrieval, Canberra distance*