ISSN: 0976-3104
ISSUE: Innovations and Economics

ARTICLE

ESTIMATION OF THE INFLUENCE OF DEMOGRAPHIC INDICATORS ON DEVELOPMENT OF REGIONAL ECONOMY USING ARTIFICIAL NEURAL NETWORKS

Vitaly Mikhailovich Tatyankin*, Vadim Faruarovic Islamutdinov

Yugra State University, Chekhova Street, 16, Khanty-Mansiysk, 628012, RUSSIA

ABSTRACT

Background: *This article describes the use of artificial neural networks for analysis of quantitative and qualitative interrelations between demographic and economic systems of northern resource-extracting region in the Russian Federation. This work is aimed at determination of demographic indicators having key influence on development of economic sectors of northern resource-extracting region in the Russian Federation.* ***Methods:*** *Methods of machine learning based on artificial neural networks were used in this work. A deep-belief neural network with sigmoid activation function was considered. The number of processing layers and neurons in them was selected according to preliminary analysis of learning sample. The learning sample and reference values were normalized with respect to the range from 0 to 2/3. The learning sample was based on data characterizing demographic indicators of the region. The indicators of economic sectors of the region were used as reference values.* ***Results:*** *An algorithm of econometric studies based on machine learning including development of mathematical model of the considered object as well as interpretation of the obtained results, was proposed. Coefficients of importance and coefficients of influence of demographic indicators on performances of development of economic sectors were obtained. Conditional correlation coefficient was determined. The obtained mathematical model of economic development of a northern resource-extracting region as a function of dynamics of demographic indicators was characterized by high degree of adequacy, which was confirmed by root mean square error of simulation equaling to 0.069 %.* ***Conclusion:*** *Assessment and verification of the proposed algorithm confirmed its high capabilities to reveal complicated interrelations between demographic indicators and performances of economic sectors of the region. The indicators of dynamics of birth rate and average income of inhabitants were the most important for development of economic sectors of a northern resource-extracting region in terms of artificial intelligence.*

KEY WORDS
Machine learning; deep-belief neural networks; demography; economic sectors; northern resource-extracting region

Received: 5 Nov 2020
Accepted: 8 Dec 2020
Published: 14 Dec 2020

***Corresponding Author**
Email:
tatyankin.v.m@mail.ru

INTRODUCTION

Social and economic systems are developed in unbreakable bond between economic and social constituents of economy. The social constituent manifests itself in statistics in the form of demographic indicators: birth rate, mortality rate, expectancy of life, etc. Therefore, it is very important to determine the influence of demography on development of a northern resource-extracting region.

At present, there are few scientific studies devoted to this topic. The work by Abdyusheva [1] should be highlighted, where regression models are proposed for birth rate and expectancy of life as a function of economic indicators. Belyaevskii [2] describes the approaches to analysis of the impact of market processes on demography and reverse impact. The mathematical apparatus is also comprised of regression analysis. In the work by Molchanova and Kruchek [3], the regression models of medical and demographic indicators are presented as a function of economic indicators. It is demonstrated that average expectancy of life directly depends on economic development. Procedural foundations of mathematical simulation of interrelation between demographic and economic indicators are considered in the thesis by Ketova [4].

Similar situation is observed in foreign publications. The works devoted to this topic are not numerous [5, 6]. In general, they state the benefits of demography development for overall national economy, however, they do not attempt to estimate qualitative and quantitative constituent of each demographic indicator.
The performed analysis of the investigated topic has demonstrated that simulation of interrelations between demographic and economic indicators is narrowly focused and applied either to single regions or to a whole country. It should be mentioned that all proposed approaches are based on regression analysis. Using this approach, it is possible to analyze the interrelation between indicators of development of economic sectors and demographic indicators of a northern resource-extracting region. However, this approach has its limitations and disadvantages. For instance, the use of standard statistic methods does not allow to reveal the implicit logically unexplainable interrelations. The implicit interrelations are revealed well by more intelligent methods, for instance, neural networks of deep learning.

In addition, the use of correlation regression analysis requires high purity and confidence of initial data, where as in the case of neural networks, errors and missed data are not so critical.

Artificial neural networks as an independent theoretical field were formed in 1950–1960 due to the works by McCulloch, Pitts, Wiener, and Rosenblatt [7–9]. Practical interest to neural networks was attracted in late 1990-s, when the backpropagation method was significantly improved simultaneously by two groups of scientists [10–11]. Almost all learning algorithms of artificial neural networks are based on this method. In 2006 G. Hinton et al. [12] proposed algorithms of deep learning of multilayer neural networks, which initiated a new stage of development of artificial neural networks. According to the scientists of Massachusetts Technological Institute, the proposed algorithms are included in the list of ten most challenging technologies capable to modify human life [13].

At present, artificial neural networks are used for solution to the following problems: image recognition, decision making, clusterization, approximation, data compression, and content-addressable memory [14]. Despite the fact that no separate trend of application of artificial neural networks in econometrics is highlighted, there are some popular works devoted to this subject [15-21].

The considered aspects confirmed high generalizing capabilities of artificial neural networks, especially upon solution to hardly formalized problems. A disadvantage of the considered solutions was that they were narrowly focused and did not allow solving the problems formulated in this work. Therefore, the algorithm has been developed to estimate the influence of demographic indicators on the indicators of development of economic sectors using artificial neural networks (hereinafter: the algorithm).

METHODS

Algorithm

The developed algorithm includes the following: description of the set of input data for econometric analysis, algorithm of the use of artificial neural networks for determination of mathematical model of the considered object, as well as interpretation of the obtained results. The algorithm assumes consecutive execution of the following actions:

i. Preparation of primary data for analysis. They can be borrowed from various databases of international statistics [22].
ii. Formation of learning sampling and references from collected primary information in Issue 1. During formation of learning sampling and references the data should be normalized with regard to the range from 0 to 0.5.
iii. Selection of neural network architecture. Deep-belief neural network is proposed as the basic type, where autoencoder method is used for pre-learning. Sigmoid should be used as activation function. In order to determine neural network architecture, the recommendations in [23] should be used.
iv. Learning of neural network to predefined error.
v. Analysis of the obtained results. This is aided by the three-step procedure:

Single signal for each neuron of input layer is passed separately through the trained neural network obtained in Issue iv. This results in the importance matrix of the indicators:

$$Importance\_matrix_{ij}, \qquad (1)$$

where i is the number of neurons in output layer of the neural network, j is the number of elements in input layer of the neural network. Using the importance matrix, we arrange the importance matrix of the considered indicators using the following equations:

$$[a_i, b_i] = \max\left(Importance\_matrix_i\right), \quad \text{for each } i \qquad (2)$$

$$IM\_Investigated\_Indicators_{b_i 2} = +1, \quad \text{for each } i \qquad (3)$$

$$[d_i, e_i] = \min\left(Importance\_matrix_i\right), \quad \text{for each } i \qquad (4)$$

$$IM\_Investigated\_Indicators_{e_i 1} = +1, \quad \text{for each } i \qquad (5)$$

where a_i is the maximum value of the string i of the matrix Importance_matrix, b_i is the number of column with the maximum a_i in the string i of the matrix Importance_matrix, d_i is the minimum value of the string i of the matrix Importance_matrix, e_i is the number of column with the minimum d_i in the string i of the matrix Importance_matrix, IM_Investigated_Indicators is the importance matrix of the considered indicators.

- The number of negative and positive weight coefficients of the input layer of trained neural network is calculated. In terms of importance, the indicators are ranked in ascending order of weight coefficient of input layer. This is a consequence of the use of sigmoidal function.

- The conditional correlation coefficient is calculated. This indicator allows to estimate the obtained regularities by common interval of correlation from -1 to 1. It is calculated from the assumption that the influence coefficients are distributed over sigmoid for which the value of 0.5 corresponds to the absence of influence (the consequence of normalization with regard to 0.5). The following equation is used for the calculations:

IIOAB JOURNAL ISSN: 0976-3104

$$CCC = (\mathrm{CI} - 0.5) * 2, \quad (6)$$

where CCC is the conditional correlation coefficient, CI is the coefficient of influence. Respectively, the coefficient of influence below 0.5 corresponds to negative values of conditional correlation coefficient.

Description of input data

In order to assess and to verify the developed algorithm, the problem of determination of interrelation between indicators of development of economic sectors and demographic indicators of northern resource-extracting region was solved using the example of Khanty–Mansi Autonomous Okrug–Yugra (KhMAO–Yugra). The following main economic sectors were considered:

- Mineral extraction.
- Processing companies.
- Wholesale and retail companies, including small business.
- Construction and production of construction materials.
- Transport.
- Financial activity.
- Agriculture, fishery, fish farming, hunting and forestry.
- Information and communication technologies.
- Production and distribution of electric energy, gas, and water.
- Service sphere, including such types of economic activity as provision of other public, social, and personal services; health care and social services; education; state management and provision of military security; real estate operations, leasing and service rendering; hotels and restaurants.

The following indicators were considered for each economic sector:

- Value added of the sector.
- Number of companies in the sector.
- Number of employees in the sector.
- Fixed assets in the sector.
- Deterioration rate of the fixed assets in the sector.
- Turnover of the companies in the sector.
- Number of shipped products in the sector.
- Import of the sector.
- Export of the sector.
- Profit and loss in the sector.
- Specific weight of unprofitable companies in the sector.
- Profitability of products in the sector.
- Investments into the sector.

Example of input data

In total 96 indicators for all sectors were derived from 2005 to 2015. Not all economic sectors were characterized by all aforementioned indicators. The initial data for mineral extraction sector are exemplified in [Table 1]. All numerical indicators of the economic sectors are summarized at GitHub [24].

Table 1: Indicators of mineral extraction industry

Indicators	Years										
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Value added, RUB bln	1,048	1,151	1,201	1,267	1,149	1,242	1,635	1,835	1,823	1,870	1,978
Number of companies	323	380	432	449	621	599	744	830	855	916	869
Number of employees, thousand persons	150.8	161.2	161.6	170.4	169	195	194.1	197.7	200.3	204.2	212.3
Fixed assets value, RUB bln	1,255	1,497	1,879	2,209	2,657	2,970	3,339	3,945	4,681	4,646	5,153
Deterioration rate of fixed assets, %	58.1	59.2	57.6	54.8	56.6	56.5	57.9	59	62	65	65.7
Turnover of the sectoral companies, RUB bln	1,285.1	1,486.2	1,539	1,831.2	1,725.1	1,904.5	2,428.5	2,767.9	2,676	2,785	3,136
Shipped amount, RUB bln	3,062	3,721	4,489	5,272	5,091	6,218	8,020	8,950	9,214	9,691	11,170
Import, USD mln	393.8	446.9	570	809.1	1,285.6	946.4	1,335.4	976	1,071.4	1,227.9	801.5
Export, USD mln	10.2	10.7	5.6	8.5	179.4	195.9	17.8	83	64.9	160.4	521.1
Profit, RUB bln	288	660	377	366	393	353	559	520	644	1,318	1,229
Specific weight of unprofitable companies, %	25	23.9	12	29.4	31.3	30.9	34.5	28.8	33.1	45.2	36.2

Product profitability, %	43.4	34	32.7	20.7	28.6	23.2	24.6	24.5	20.5	16.5	21.1
Investments, RUB bln	140	205	259	320	314	351	459	462	478	510	576

The following demographic indicators were used [Table 2]:

- Population, birth rate, per 10 thousand persons;
- Birth rate, per 10 thousand persons;
- Mortality rate, per 10 thousand persons;
- Expectancy of life, years;
- Marriage rate, per 10 thousand persons;
- Divorce rate, per 10 thousand persons;
- Migration gain, per 10 thousand persons;
- Average income of inhabitants, RUB.

The data source was the Federal State Statistics Service [25].

Table 2: Demographic indicators of KhMAO–Yugra

Indicators	Years										
	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
1	1,468	1,488	1,505	1,520	1,539	1,537	1,561	1,584	1,597	1,612	1,626
2	13.6	13.7	14.6	15.3	15.6	16.4	16.4	17.7	17.5	17.2	16.6
3	7.1	6.8	6.7	6.8	6.6	6.8	6.5	6.3	6.3	6.4	6.4
4	67.82	68.84	69.35	69.91	70.45	70.30	70.91	71.79	72.23	72.27	72.58
5	9.7	10.5	11.5	10.6	10.7	11.3	12	10.6	10.5	10	9.3
6	6.7	6.9	7	7.1	7	7.1	7.1	6.6	6.3	6.3	5.7
7	-27	-1	34	11	32	8	57	32	-29	-16	-11
8	18,115	22,380	27,110	32,872	32,263	32,385	33,926	36,345	39,292	41,503	44,538

RESULTS AND DISCUSSION

The numerical assessment of the proposed method using the data of KhMAO–Yugra resulted in deep-belief neural network with the learning error of 0.069 %.

[Table 3] summarizes the results for the importance matrix of the considered indicators. The weight coefficients of input layer of trained deep-belief neural network are analyzed. Detailed analysis of the obtained numerical results can be found at GitHub [24].

Table 3: Importance matrix of the considered indicators

Indicators	Less important	Most important	Difference
Population	0	2	2
Birth rate	21	42	21
Death rate	39	13	-26
Expectancy of life	6	7	1
Marriage rate	9	11	2
Divorce rate	9	5	-4
Migration gain	9	2	-7
Average income of inhabitants	3	14	11

According to [Table 3], the birth rate is the most important indicator for development of sectors in KhMAO–Yugra, the average income of inhabitants occupies the second position. The indicators of the least influence are the mortality rate, the migration gain, and the divorce rate. The influence of such indicators as the expectancy of life, the marriage rate, and the population is weak, if any. The migration gain should be considered separately. Its weak influence on economic development is stipulated mainly by the fact that this indicator is negative, that is, higher number of persons leave the Okrug in comparison with newcomers. Moreover, this is highly volatile indicator, which influences the simulation quality.

Analysis of [Table 4] confirmed the trends revealed upon consideration of [Table 3]. The birth rate is again the most important indicator for development of economy sectors. The second position is shared by the average income of inhabitants and the marriage rate. The mortality rate also leads by a wide margin with respect to negative influence on development of economy sectors.

IIOAB JOURNAL
ISSN: 0976-3104

Table 4: Statistics of input layer of trained deep-belief neural network

Indicators	Negative dependence	Positive dependence	Average weight coefficients
Population	8	7	-0.32
Birth rate	11	4	-4.91
Death rate	9	6	0.71
Expectancy of life	8	7	-0.05
Marriage rate	8	7	-2.50
Divorce rate	9	6	-1.65
Migration gain	9	6	-1.01
Average income of inhabitants	7	8	-3.10

Conditional correlation coefficient

[Tables 5–8] summarize the conditional correlation coefficients for the birth rate and the average income of inhabitants.

Table 5: Conditional correlation coefficient for birth rate. Indicators of sectors with positive influence

Indicators of sectors	Coefficient of influence	Conditional correlation coefficient
Specific weight of unprofitable agricultural companies	0.999896	+0.999792
Product profitability of power engineering companies	0.971219	+0.942438
Fixed assets of mineral extraction companies	0.961313	+0.922626
Turnover of construction companies	0.958428	+0.916856
Turnover of wholesale and retail companies in KhMAO–Yugra	0.953416	+0.906832

Analysis of [Table 5] demonstrated significant direct influence of the birth rate on the specific weight of unprofitable agricultural companies, the product profitability of power engineering companies, and the fixed assets of mineral extraction companies. However, this is attributed to certain random factors, whereas the interrelation of the birth rate with such indicators as the turnover of construction companies and the turnover of wholesale and retail companies in KhMAO–Yugra is attributed to increase in consumer demands in the considered sectors.

Table 6: Conditional correlation coefficient for birth rate. Indicators of sectors with negative influence

Indicators of sectors	Coefficient of influence	Conditional correlation coefficient
Passengers transported by motor vehicles	0.0590	-0.881
Export, USD mln	0.0629	-0.874
Profit, loss in mineral extraction sector	0.1409	-0.718

In [Table 6] the interrelation of birth rate with the number of transported passengers and financial results of mineral extraction sector is random, whereas the negative dependence of export can be attributed to increased consumption in the region.

Table 7: Conditional correlation coefficient for average income of inhabitants. Indicators of sectors with positive influence

Indicators of sectors	Coefficient of influence	Conditional correlation coefficient
Shipped products in construction industry	0.999972	+0.999944
Fixed assets of processing companies in KhMAO–Yugra	0.996754	+0.993508
Balanced financial result (profit, minus, loss) for economy in total	0.996221	+0.992442
Fixed assets of construction sector	0.995303	+0.990606
Cost of products of processing companies in KhMAO–Yugra	0.992645	+0.98529

In [Table 7] the positive influence of average income of inhabitants on balanced financial results could be probably attributed to coincidence of income dynamics of inhabitants and companies. And the positive influence of average income of inhabitants on the indicators of construction and processing companies is

IIOAB JOURNAL ISSN: 0976-3104

quite obvious, since the income increase results in increase in consumption. Perhaps, in the region this is expressed in increased expenses for purchase and construction of housing as well as products of processing companies.

Table 8: Conditional correlation coefficient for average income of inhabitants. Indicators of sectors with negative influence

Indicators of sectors	Coefficient of influence	Conditional correlation coefficient
Profit and loss in transportation sector	0.000878	-0.998244
Norm of federal spending (taxation)	0.071524	-0.856952
Amount of freight transport by road	0.116699	-0.766602

Negative influence of average income of inhabitants on indicators of transportation sector in [Table 8] can be attributed to the fact that with the increase in incomes, the structure of consumption of transportation services varies in favor of personal vehicles. The negative influence on federal spending is attributed to the fact that the region is a budget donor, high portion of rent is withdrawn in the form of taxes (about 90%). Therefore, the increase in the incomes of inhabitants in the region is possible only with the decrease in tax burden, when more money is retained in the region.

Novelty of the research

The analysis of the subject area has shown that there are no works devoted to the definition of demographic indicators that have key impact on the development of economic sectors in the northern resource-extracting region of the Russian Federation. In order to solve this problem, an econometric research algorithm based on machine learning has been developed, which allowed identifying implicit, logically inexplicable relationships. The peculiarity of the algorithm consists in determining the coefficients of importance and the coefficients of demographic indicators influence on the indicators of the development of branches of the regional economy using the parameters of the trained neural network.

CONCLUSION

An algorithm for econometric research based on machine learning has been proposed, including the construction of a mathematical model of the object under study, as well as the interpretation of the results obtained. Approbation and multiple verification of the proposed algorithm have confirmed its high ability to identify complex patterns between demographic indicators and indicators of sectors of the regional economy. The mathematical model of development of sectors of a northern resource-extracting region as a function of dynamics of demographic indicators, developed by means of deep-belief neural network, is characterized by high rate of adequacy. This is confirmed by the root mean square error of simulation equaling to 0.069%. The coefficients of importance and coefficients of demographic indicators influence on the indicators of development of the regional economy branches have been obtained. The most important for the development of economic sectors of the northern resource-extracting region from the point of view of artificial intelligence are indicators of the dynamics of "birth rate" and "average income of inhabitants." At the same time, the mortality rate is by far the leader in terms of negative impact on the development of economic sectors. Conditional coefficients of correlation between the "birth rate" and "average income of inhabitants" indicators for the sectors of the economy have been calculated. Further development of the obtained method consists in the expansion of geography of studies by addition of new regions. This would improve the simulation quality as a consequence of building up of primary data for analysis. Such increase in the data amount and, respectively, forecast quality can be achieved by using data for shorter time intervals: quarter, month, if they are available. In addition, a very challenging trend is involvement of information from big databases as well as free online data into scientific turnover. Separate investigation is planned regarding the influence of demographic indicators on the properties of development of regional economic sectors using standard econometric approaches.

CONFLICT OF INTEREST
There is no conflict of interest.

ACKNOWLEDGEMENTS
None.

FINANCIAL DISCLOSURE
None.

REFERENCES#

[1] Abdyusheva SR. [2015] Estimation of interrelation between demographic and economic indicators. Proceedings of 4th All-Russian Conference Devoted to the 75th Anniversary of Faculty of Physics and Mathematics. Ufa, RU.

[2] Belyavskii IK. [2014] Social demographic marketing: problems, objectives, analysis. Ekonomicheskaya teoriya, analiz, praktika, 3:92–111.

[3] Molchanova EV, Kruchek MM. [2015] Estimation of interrelations of medical and demographic with social and

economic indicators of regional developments using economic and mathematical methods. Proceedings of 6th Ural Demographic Forum: Demographic Processes in Post-Soviet Territory, RU
[4] Ketova KV. [2004] Optimum distribution of capital investments with consideration for demographic forecast: cand. thesis. Izhevsk.
[5] Lucas RE. [1988] On the mechanics of economic development. Journal of Monetary Economics, 22:3–42.
[6] Schultz ThW. [1981] Investing in People. The Economics of Population Quality. Berkeley: University of California Press, USA.
[7] Wiener N. [1961] Cybernetics: Or Control and Communication in the Animal and the Machine. Paris: Hermann & Cie, Camb. Mass. (MIT Press), USA.
[8] Rosenblatt F. [1986] Principles of neurodynamics; perceptrons and the theory of brain mechanisms. In: Palm G, Aertsen A (eds) Brain Theory. Springer, Berlin, Heidelberg. DOI: 10.1007/978-3-642-70911-1_20.
[9] McCulloch WS, Pitts W. [1943] A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys, 5:115–133.
[10] Rumelhart DE, Hinton GE, Williams RJ. [1986] Learning Internal Representations by Error Propagation, in: Parallel Distributed Processing, Cambridge: MIT Press, USA, p318 – 362.
[11] Bartsev SI, Okhonin VA. [1986] Adaptive data processing networks. Krasnoyarsk: Institute of Physics, Academy of Sciences of the USSR, Siberian Branch, RU
[12] Hinton GE, Osindero S, Teh YW. [2006] A fast learning algorithm for deep belief nets. Neural Computation, 18(7):1527–1554.
[13] Golovko V. [2015] From multilayers perceptrons to deep belief neural networks: Training paradigms and application, Lections on Neuroinformatics, Golovko, VA. (Edt.), Moscow: NRNU MEPhIp, p47–84.
[14] Artificial neural network. Retrieved from: https://en.wikipedia.org/wiki/Artificial_neural_network
[15] Kuan CM, White H. [1994] Artificial neural networks: An econometric perspective. Econometric Reviews, 13(1):1–91.
[16] Luxhøj JT, Riis JO, Stensballe B. [1996] A hybrid econometric-neural network modeling approach for sales forecasting. International Journal of Production Economics, 43(2–3):175–192.
[17] Hruschka H. [1993] Determining market response functions by neural network modeling: A comparison to econometric techniques. European Journal of Operational Research, 66(1):27–35.
[18] Moshiri S, Cameron N. [2000] Neural network versus econometric models in forecasting inflation. Journal of Forecasting, 19(3):201–217.
[19] Church KB, Curram SP. [1996] Forecasting consumers' expenditure: A comparison between econometric and neural network models. International Journal of Forecasting, 12(2):255–267.
[20] Frey UJ, Rusch H. [2013]. Using Artificial Neural Networks for the Analysis of Social-Ecological Systems. Ecology and Society, 18(2):40.
[21] Mohammadfam I, Soltanzadeh A, Moghimbeigi A, Savareh BA. [2015] Use of Artificial Neural Networks (ANNs) for the Analysis and Modeling of Factors That Affect Occupational Injuries in Large Construction Industries. Electron Physician, 7(7):1515–1522.
[22] Databases of international statistics. Retrieved from: http://lib.isert-ran.ru/files/lib/indexes_and_stats.pdf
[23] Golovko VA, Krasnoproshin VV. Neirosetevye tekhnologii obrabotki dannykh [Neural Network Data Processing Technologies]. Minsk BGU.
[24] GitHub, Inc. Retrieved from: https://github.com/VitalyHMAO/Data
[25] Federal State Statistics Service, Statistical reports. Retrieved from: http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/publications/catalog

#*English translations of the references are presented.*