Neural network modelling for determining the priority areas of regional development

A K Moskalev, A E Petrunina, N S Tsygankov

Siberian Federal University, 28, Kirensky St., Krasnoyarsk 660074, Russia

Email: ak_moskalev@mail.ru

Abstract. Neural network modeling based on self-organizing Kohonen maps was carried out in order to cluster the economy of one of the administrative districts of the Krasnoyarsk Territory. The research is based on statistical data on the economy of about one hundred enterprises of various spheres operating in the region. It is shown that in the process of clustering four groups of enterprises can be clearly distinguished according to the types of economic activity. Having applied the neural network modelling method, we identified three enterprises within the framework of this cluster, which can be considered as “growth areas” of the district economy. Self-organizing maps that is trained using unsupervised learning make it possible to get an idea not only of promising areas of development, but also to identify key parameters that ensure the leadership and competitiveness of the region.

1. Introduction

The evolution of the cluster theory in economics has happened for several decades already, but this topic still does not lose its relevance [1]. Today in Russia, cluster approach is the most effective method of development of the regions, since as a result of the functional integration, the businesses have not only increased their profitability, but also, due to the active cluster interactions, improved and multiplied competitive advantages for both cluster members and the region as a whole [2].

Cluster formations, being independent and self-organized economic units, are able to ensure not only quantitative growth of the economy, but also to act as stimulators of innovative activity in the region and certain interrelated industries [2-4]. The cluster approach as a technology of regional development management and, accordingly, the cluster itself as an element of innovation infrastructure, create opportunities for the region and businesses not only to survive, but also to develop during the transition to an innovative model of development [5-6].

There are different approaches to the definition of a cluster. Analyzing them taken together, we can say that they have one common feature, namely the integration of any objects that are allocated from a large population on the grounds of a certain identical group characteristic. Emphasizing the distinctive features of clusters in each of the sciences, one shall note that the universal features of clusters can be the following: homogeneity of components for the mutual efficient exchange of available resources; the combination of centripetal and centrifugal forces; the fuzziness and mobility of boundaries and configuration; the shared resource base.

The following methods are used to determine cluster specialization [7-9]:
- analytical method: the study of data on the territory, its natural, labour and financial resources of the productive forces.
- method of indicative planning of competitiveness: quantitative and qualitative analyses of the present competitiveness and determination of the prospective one.
- modelling method: the study of cluster objects by building and studying models of real organizations, processes or phenomena.
- the “input-output” method: the study of the circulation of goods and services between industries and other sectors of the economy.

2. Methods
In this study, we used a neural network method to model the specialization of the cluster, and Kohonen self-organizing maps as a clustering tool. The simulation was carried out using special software – analytical platform Deductor Studio to build a neural network trained by unsupervised learning method based on the structure of the input data.

The classification method based on self-organizing Kohonen maps is as follows:
- determination of the input data vector based on the results of the training experiment;
- determination of parameters for building a network;
- analysis of the results with finding an estimate of the classification accuracy.

To organize a network, you must define the following construction parameters:
- the size of the training and test set;
- network size (number of cells);
- an example recognition criterion (the value of the recognition error);
- learning speed and radius;
- the number of clusters.

Since the sample size is small, the training set should be as large as possible and sufficient for the correct organization of the network. In this case, the test set should be sufficient for testing the network with 100% traffic. Based on the stated conditions, the size of the training set was 78 copies (91% of the sample), the size of the test set was 8 copies (9% of the sample).

The size of the network was chosen in such a way as to exclude the appearance of “dead” neurons; such neurons, the vector of synaptic weights of which is significantly removed from the vector of input data. Such neurons are not used when training the network, because cannot win competition from nearby neurons. This situation leads to the fact that the input data is interpreted by a smaller number of neurons than was originally specified, which introduces distortions in the final result. To study the presented samples, networks with a dimension of 16x12 cells were selected.

The type of cells is selected as hexagonal: it more correctly displays the Cartesian distance between the centers of the cells. The rest of the parameters for constructing the Kohonen network are presented in table 1.

| Network parameter | Specified Criterion |
|-------------------|---------------------|
| Recognition error rate | An example is considered recognized if the recognition error is less than 0,05% |
| Number of eras | 250 |
| Learning speed | At the beginning of training: 0,4 At the end of the training: 0,005 |
| Training radius | At the beginning of training: 2 At the end of the training: 0,2 |
| Clustering | Fixed number of clusters: 4 [6] |

Kansky district of the Krasnoyarsk Krai was chosen as an object of clustering for several reasons:
- need for the development of territories, including districts of the Krasnoyarsk Krai, is determined by the strategy of social and economic development of Siberia until 2020;
- agriculture and forestry sectors prevailing in the region are designated as priority areas in this programme;
- low indicators of efficiency of economy of the area in comparison with nearby territories of Krasnoyarsk Krai (Nazarovsky, Achinskiy, Sharypovsky Districts).

The analysis of the economy made on the basis of the study of the economic activities of the leading enterprises of the Kansky District showed that there are four main areas of activity:
- meat and dairy production;
- plant growing;
- forestry;
- other companies (commodity and resource-providing producers).  
The database contains economic indicators of 86 enterprises of the Kansky District with different spheres  
of activity, including:  
- 12 enterprises producing meat and dairy products;  
- 15 enterprises producing crop products;  
- 46 enterprises in the sphere of forestry;  
- 13 resource-providing enterprises.  
Kansk, the administrative centre of Kansky District, is considered to be the core of the future cluster.  
During the simulation, the following parameters were selected as input:  
- cost of sales (RUB);  
- net profit (RUB);  
- distance from the cluster core (km);  
- number of employees (people);  
- type of activity.  

3. Results  
The result of modelling self-organizing maps is presented in Figure 1. A fairly homogeneous structure is  
visible, which corresponds to the sectoral affiliation of the District’s economy.

![Figure 1. Visualization of clustering results. Add legend and describe the process.](image-url)

The X axis shows the correspondence of the attribute value and the color gradient depending on its value.  
Cells in the figure characterize the number of neurons used in the neural network (192). Boundaries, in black, characterize the selection of values that fall into one cluster. The top left card is the Distance from  
the cluster core indicator, the top right card is the Number of employees indicator, the middle left card is  
the Cost of sales indicator, the middle right card is the Net profit indicator, the lower left card is the indicator "Type of activity", bottom right map - general map of the distribution of the resulting neurons  
into clusters.  
Neural network modelling showed the presence of four distinct clusters, the sizes of which differ slightly.  
Taking into account the topological similarity of companies, the indicator of the type of activity was  
singled out as the main one.  
Figure 2 shows the significance of the indicators for creating modelling maps. The first cluster combines  
companies producing meat and dairy products, and processing companies with the highest  
competitiveness. Also, the results of modelling show that the companies of this sector have a similar number of staff and the cost of production, so these indicators are also set as having a high importance.
The second cluster includes companies producing crop products and companies of the processing industry. The indicators that determine the specifics of the companies of this cluster include the type of activity, the number of staff and the annual net profit of the company.

Table 2 - Cluster profile

| Field                        | Indicator | Clusters          |
|------------------------------|-----------|-------------------|
|                              |           | 1          | 2           | 3          | 4          | TOTAL   |
| Number of enterprises in cluster | Number   | 46 (53.5%) | 17 (19.8%) | 15 (17.4%) | 8 (9.3%)  | 86 (100%) |
| Type of activity             | Significance | 100.0%      | 100.0%      | 50.3%      | 93.8%      | 100.0%  |
| Number of employees          | Significance | 97.1%      | 94.2%      | 82.1%      | 99.8%      | 100.0%  |
| Net profit                   | Significance | 93.6%      | 72.8%      | 63.4%      | 98.7%      | 99.72%  |
| Cost of sales                | Significance | 93.8%      | 93.2%      | 15.5%      | 64.8%      | 97.42%  |
| Distance from the cluster core | Significance | 86.8%      | 43.4%      | 73.7%      | 78.7%      | 91.62%  |

The third cluster includes resource-providing companies. These companies provide:
- water supply, power supply and heat supply;
- equipment supply.
In this cluster, the type of activity is no longer decisive, as the key ones are the distance from the core of the cluster and the number of employees.
In the fourth cluster, companies related to forestry were combined. This cluster is formed by the largest number of companies, so this direction is the main candidate for leadership.
However, analyzing the requirements and recommendations for the structure of cluster formations [2], it can be concluded that in addition to industry companies, the clusters should include:
- suppliers of raw materials and equipment;
- financial institution;
- research institutes;
- transport and trade intermediaries;
- waste processing and recycling facilities;
- social infrastructure companies.
On this occasion, it can be concluded that the agricultural and industrial cluster is the most promising tendency of economic development of the Kansky District.
Based on the received information and the requirements for companies that should join the cluster, we may offer to introduce into the agricultural cluster of Kansky District the companies from the first group, which form the core of a cluster, companies from the second group, which supplement a system of agricultural products, as well as some companies of the third group, supporting the activities of major companies.
It is imperative to note that the indicator of territorial remoteness from the centre of the cluster was not identified as determining what is adequate in modern conditions of digital transformation and widespread development of communication technologies.
Based on the requirements for the companies that must be included to the cluster we find it necessary to add the following companies to the above mentioned ones:
- veterinary institutions (3 organizations in the District);
- scientific and educational institutions (more than 10 institutions in the District);
- banking institutions (branches of more than 15 different banks).
In the cluster formed during the modelling, it is possible to identify key “growth areas”, i.e. companies that are mostly important for this cluster. For this purpose, using statistical data only of the first group companies, a modelling procedure was carried out, the results are presented in figure 2.
The X axis shows the correspondence of the attribute value and the color gradient depending on its value. Cells characterize the number of neurons used in the neural network (192). Boundaries, in black, characterize the selection of values that fall into one cluster. Top left map - Distance from the cluster core, upper right map - Number of employees, left middle map - Cost of sales, right middle map - Net profit, bottom map - general distribution map final neurons per clusters.

4. Discussion of the results
As a result of modelling, three companies of the Kansky District with the criterion of maximum net profit were selected, and they can be the key/flagship enterprises of the cluster. This research allows us to conclude that the use of the neural network approach in the formation of the list of priority areas of economic development of regions has great prospects. Self-organizing maps that support the algorithm of teaching without a teacher, enable one to get an idea not only about the promising directions of development, but also to identify the crucial parameters that ensure the leadership and competitiveness of the region.

The application of the considered approach of using a neural network for clustering enterprises within a region can be used in the development of comprehensive programs for regional innovative development, as well as in the development of a model for the functioning of a regional innovation infrastructure in terms of methods for selecting enterprises and projects to support individual objects of innovation infrastructure.

Acknowledgments
The reported study was funded by RFBR, project number 19-37-90023.

References
[1]. Markov L S 2015 Theoretical and methodological foundations of the cluster approach (Novosibirsk, SB RAS Publ.) 300
[2]. Abramov R A, Podchufarov S K 2017 Regional innovation potential of industrial clusters. Reg. Econ.: Theory and Practice 15 5 881-895
[3]. Kotliarov N N 2014 Foreign experience in the formation of cluster systems World Econ. and Int. Econ. Relations 10 105-110
[4]. Tonkova S, Burmatova O P 2016 Bulgaria’s experience in the formation of cluster policy Region: Econ. and Sociology 2 237-252
[5]. Burmatova O P 2016 Clusters as a form of local production systems Econ. and Management World 16 4 93-108
[6]. Zemtsov S P, Bukov D V 2016 Methods for identifying clusters of small and medium-sized businesses *Regional Econ.: Theory and Practice* 3 104-117

[7]. Glukhov, V. V., & Balashova, E. (2015). Operations strategies in info-communication companies. doi:10.1007/978-3-319-23126-6_48

[8]. Ivanov, D., Sokolov, B., Serova, E., & Yusupov, R. (2020). Combined approach to the complex objects control and stability analysis of management decisions. *International Journal of Risk Assessment and Management*, 23(1), 106-118. doi:10.1504/IJRAME.2020.106175

[9]. Shestak, O., Shcheka, O. L., & Klochkov, Y. (2020). Methodological aspects of use of countries experience in determining the directions of the strategic development of the russian federation arctic regions. *International Journal of Systems Assurance Engineering and Management*, 11, 44-62. doi:10.1007/s13198-019-00805-w