

*Research article*DOI: <https://doi.org/10.48554/SDEE.2021.1.7>

CLUSTERING OF TERRITORIAL OBJECTS IN THE MANAGEMENT OF THEIR SUSTAINABLE DEVELOPMENT

Dmitrii Rodionov¹ , Dmitrii Alferyev^{1,2*} , Yuliya Klimova² , Kaisar Alpysbayev³ ¹ Peter the Great St. Petersburg Polytechnic University, Russia, drodionov@spbstu.ru, alferev_1991@mail.ru² Vologda Scientific Center of the Russian Academy of Sciences, Russia, j.uschakowa2017@yandex.ru³ International Information Technology University JSC, Almaty, Kazakhstan, kaisaralp@gmail.com* Corresponding author: alferev_1991@mail.ru

Abstract

It is necessary to understand the nature of spatial territorial socio-economic objects in order for them to have an effective influence on the implementation of measures intended to increase the living standards of the population that resides there. To achieve this, they must be correctly identified amongst a general set of objects. In this regard, the purpose of this work is to develop a tool for territorial clustering. Science is one of the engines of socio-economic progress through which innovations are implemented. Hence, we test the clustering of territorial objects (regions of Russia) in relation to statistical financial cost data for science in terms of their relationship with wages and incomes of the population, the GRP (Gross Regional Product) and innovation activity. The main tool used for cluster analysis is the perceptron mathematical model, the features of which we describe in detail in this work. It follows from its characteristic features that it divides a studied population in a manner that allows for the possibility to simulate the increasing or decreasing dynamics of one quantity's dependence on another. The study develops a universal algorithm for the purpose of territorial cluster analysis, which is proven in the construction of the final models of dependence (paired linear regression) of the indicators identified in the work, whose coefficient of determination is primarily 0.8. In our conclusion, we indicate possible options for the further development of this study, both with respect to the technical aspects of refining and improving the algorithm as well as within the framework of a more detailed analysis of the identified regression patterns using the example of statistical data of Russian reality in relation to science and the level of life quality.

Keywords: Clustering, perceptron, spatial economics, modelling of economic processes, econometric analysis, science and innovation.

Citation: Rodionov, D., Alferyev, D., Klimova, Yu., Alpysbayev, K. (2021). Clustering of territorial objects in the management of their sustainable development. *Sustainable Development and Engineering Economics* 1, 7. <https://doi.org/10.48554/SDEE.2021.1.7>

This work is licensed under a [CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/)

© Rodionov, D., Alferyev, D., Klimova, Yu., Alpysbayev, K., 2021. Published by Peter the Great St. Petersburg Polytechnic University

Научная статья

УДК 332.14

DOI: <https://doi.org/10.48554/SDEE.2021.1.7>

КЛАСТЕРИЗАЦИЯ ТЕРРИТОРИАЛЬНЫХ ОБЪЕКТОВ В УПРАВЛЕНИИ ИХ УСТОЙЧИВЫМ РАЗВИТИЕМ

Дмитрий Родионов¹ , Дмитрий Алферьев^{1,2*} , Юлия Климова² , Кайсар Алпысбаев³ 

¹ Санкт-Петербургский политехнический университет Петра Великого, Россия, drodionov@spbstu.ru, alferev_1991@mail.ru

² Вологодский научный центр РАН, Россия, j.uschakowa2017@yandex.ru

³ Международный Университет Информационных Технологий, Алматы, Казахстан, kaisaralp@gmail.com

*Автор, ответственный за переписку: alferev_1991@mail.ru

Аннотация

Понимание природы пространственных территориальных социально-экономических объектов необходимо для эффективного воздействия внутри них для реализации мер по увеличению качества жизни населения, которое там проживает. Для этого среди общей совокупности объектов их необходимо правильно идентифицировать. В этой связи цель данной работы заключается в разработке инструментария кластеризации территорий. Одним из двигателей социально-экономического прогресса выступает наука, посредством которой воплощаются в жизнь инновации. На основании этого кластеризация территориальных объектов (регионов России) будет апробирована на статистических данных финансовых затрат на науку в их взаимосвязи с оплатой труда и доходами населения, ВРП, и инновационной активностью. Основным инструментом кластерного анализа определена математическая модель перцептрона, особенности которой детально описаны в работе. Из ее характерных черт следует выделить то, что она делит исследуемую совокупность таким образом, что сохраняется возможность моделирования возрастающей или снижающейся динамики зависимости одной величины от другой. Итоговым результатом исследования стала разработка универсального алгоритма кластерного анализа территорий, который подтвердил себя при построении конечных моделей зависимости (парная линейная регрессия) обозначенных в работе показателей, коэффициент детерминации которых у большинства равен 0.8. В заключении обозначены возможные варианты дальнейшего развития исследования как в направлении технических аспектов доработки и совершенствования алгоритма, так и в рамках более детального анализа по выявленным регрессионным закономерностям на примере статистических данных российской действительности в отношении науки у уровня качества жизни.

Ключевые слова: кластеризация, перцептрон, пространственная экономика, моделирование экономических процессов, эконометрический анализ, наука и инновации.

Цитирование: Родионов, Д., Алферьев, Д., Климова, Ю., Алпысбаев, К. (2021). Кластеризация территориальных объектов в управлении их устойчивым развитием. *Sustainable Development and Engineering Economics* 1, 7. <https://doi.org/10.48554/SDEE.2021.1.7>



Эта работа распространяется под лицензией [CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/)

© Родионов, Д., Алферьев, Д., Климова, Ю., Алпысбаев, К., 2021. Издатель: Санкт-Петербургский политехнический университет Петра Великого

Introduction

Advances in science and technology are the driving forces of economic and social development, affecting economic growth, product quality, population living standards and so on. This fundamental idea has been studied in detail in the works of the Austro-American economist Schumpeter (1980). The existence of such patterns is described in detail in the work of Stepanova and Lesnikova ‘The Role of Innovations in the Modern Development of Russian Society’ (2017) and in the article by Lugovaya ‘Innovations as the Basis for the Modernization of Modern Society’ (2012). Funding for research and development (R&D) plays an important role in the process of creating innovations. In 2018, the share of R&D funding costs in the gross domestic product (GDP) in the Organisation for Economic Co-operation and Development (OECD) countries was 4.5% (including Sweden: 3.3%, Austria: 3.2%, Germany: 3.1%, the UK: 1.7%, Japan: 3.3%, Korea: 4.5%, China: 2.1%). However, in Russia, the volume of R&D costs remains at an extremely low level. In the 2015–2017 period this indicator was 1.1%, decreasing to 0.98% in 2018, which is comparable to the indicators for South Africa, Brazil and Slovakia (about 1.0% of GDP).¹ One of the ways in which the problem of low R&D costs can be solved is through the creation of funds to support scientific, technical and innovation activities – an important aspect of which is the provision of financial support for R&D.

This study suggests that the creation of funds in order to support scientific, technical and innovation activities can have a significant impact on the socio-economic development of a country. To do so, it is first necessary to determine the relationship between an indicator such as ‘R&D costs’ and other parameters that characterise a population’s living standard, a country’s economic development, etc. The formation of such mathematical models of relationships would not only allow us to achieve certain desired results via inertia but would also make it possible for us to create a system of measures for them so that they remain stable over an extended period.

When determining such relationships within seemingly identical territorial objects, a problem arises because similar processes and phenomena occur in these objects in different ways. Consequently, there is a need for the studied objects to correctly be correlated into groups within which it would be possible to apply classical and proven methods of data processing and analysis.^{2,3}

Therefore, the purpose of this work is to develop a toolkit for clustering territories. Its approbation is carried out on the data associated with the assessment of the impact of investments in science on the level of the population’s well-being, characterised through the prism of various statistical metrics. The need for cluster analysis of territories in this direction is due to the identification of their priority areas of scientific research for the implementation of local administrative measures. These measures, in turn, would more quickly enable faster growth in the population’s well-being in areas in which appropriate scientific directions are implemented and specific innovative projects are developed.

2. Literature review

First, we briefly describe what positions on the issue of assessing the impact of investments in science are indicated in modern scientific literature. A literature review reveals that there are different views amongst researchers regarding what indicators affect the R&D cost amounts and, conversely,

¹ Gross domestic spending on R&D, (n.d.). <https://data.oecd.org/rd/gross-domestic-spending-on-r-d.htm>

² Ayvazyan, S.A., 2010. *Methods of Econometrics: Textbook*, Master. INFRA-M, Moscow

³ Marno, V., 2008. *Guide to Modern Econometrics*. Scientific Book, Moscow

how R&D financing affects other parameters. The study of the scientific literature has shown that, in general, different authors do not identify indicators but factors that can somehow influence the R&D financing.

For example, according to Yegorenko et al. (2018), R&D financing consists of the following components: federal budget, commercial organisations, non-profit sector and international investment. At the same time, it is important to note that, according to these authors, commercial organisations have a significant impact on the growth of R&D costs. According to OECD data, in most developed countries (China, the Republic of Korea, Japan, the United States, Germany, the United Kingdom, France, etc.), the share of the commercial sector in the country's R&D costs exceeds 40%, while in Russia this indicator is at only 28.1%. In China, for example, the state contributes only a fifth of the total R&D investment, while the business sector directs more than 76% of the funding.⁴

According to Seidl da Fonseca and Pinheiro-Velos (2018), the R&D cost amount can be influenced by such factors as the availability of venture funds that are designed to help companies at different stages of development. In addition, the possibility of obtaining any tax benefits in the field of scientific, technical and innovation activities, as well as the availability of a favourable legislative environment, can be important parameters that affect R&D financing. According to a team of authors led by Seidl de Fonseca (2018), taxes can have a serious impact along with the risks that always accompany all innovative projects.

It is important to note that many authors (Rodina, 2014; Yurchenko, 2013; etc.) emphasise tax incentives and a favourable legislative environment as some of the factors affecting the growth of R&D costs. According to Pashintseva (2018), there is a relationship not only between such indicators as R&D costs, federal budget and availability of venture funds but also between R&D funding and the net profit of organisations.

In addition, as Zhukovskaya et al. (2021) emphasise, the increase in R&D costs does not result from an increase in funding, an increase in the interest of both the state and private investors in the renewal of equipment and technologies or the involvement of R&D results in commercial turnover but from indexation to the level of inflation.

At the same time, when analysing the scientific literature, it is also found that there is a relationship between R&D financing and the foreign policy situation (Maslova and Lalaeva, 2018).

Thus, it is important to note that a significant number of authors do not name specific indicators but only highlight the presence of factors that are somehow related to R&D costs. Nevertheless, the analysis of the scientific literature allows us to identify the parameters that characterise the dependence on the amount of R&D funding, which include: federal budget, commercial and non-profit sectors, international investment, foreign policy environment, taxes, availability of tax incentive tools, favourable legislative environment, availability of venture funds, risks, GDP, inflation and so on.

However, it is important to note here that it is difficult to carry out calculations in order to assess the relationship between changes in R&D costs and the other above-mentioned parameters because many authors do not discuss specific indicators, with the exception of GDP, inflation, international investment and federal budget. Factors such as commercial and non-commercial sectors do not provide a clear understanding of what indicators are being referred to by the authors. At the same time, factors such as foreign policy environment and tax incentive instruments are generally difficult to describe statistically, making it difficult to use these parameters. Hence, it is necessary to look for additional indicators in order to find the relationships between R&D costs and other parameters.

⁴ Gross domestic expenditure on R&D by sector of performance and source of funds, (n.d.). https://stats.oecd.org/Index.aspx?DataSetCode=GERD_SOF

As mentioned above, we assume that the change in R&D costs is related to the parameters of socio-economic development. In this regard, based on the data of the Federal State Statistics Service, we propose to use wages, income of the population, GDP and innovation activity as the main indicators that characterise the population's standard of living as well as the economic and innovative development. Accordingly, it is necessary to analyse the dependence of these indicators on changes in the volume of R&D financing and vice versa.

The issue does not end here and rests on the fact that the dynamics of the above-mentioned indicators behave differently. This is due to the different spatial features of the studied territories. In a series of papers by Kudryavtseva and Skhvediani (2020a, 2020b), the authors discuss the relevance of finding solutions to such problems in detail. In the article 'Econometric Analysis of the Industry Specialization of the Region: on the example of the Manufacturing Industry of Russia' (Kudryavtseva and Skhvediani, 2020a), the author team proposes several tools for assessing regional specifics in accordance with the industrial production located on their territories. In the article 'Studying Regional Clusters with the Use of Data Processing Systems: The Case of the Biopharmaceutical Cluster' (Kudryavtseva and Skhvediani, 2020b), the authors managed to distinguish regions into separate groups in accordance with estimates of the 'localisation', 'size' and 'focus' of a biopharmaceutical cluster located in the territorial space of Russia.

The problems with assessing territorial objects, their development and functioning are also presented in a number of other Russian works. Thus, in Kozhevnikov's (2019) 'Spatial and territorial development of the European North of Russia: Trends and priorities of transformation', the author identifies problems of regional management and highlights their features for the northern areas of the Russian Federation. In Alferyev's (2018) talking points, the work of the autoregressive model is demonstrated on the basis of an example of the Republic of Belarus regions cooperation in science and technology. An article by Minakir (2017) covers developments on spatial and territorial topics in general, analysing the main achievements and developments in this area. The article by Fonotov and Bergal' (2020) provides an overview of foreign developments in the implementation of the policies of individual territorial subjects of states and clusters formed on these states.

A number of foreign works are also devoted to the topic of territorial subject clustering and of the resulting administrative impact on them. Ketels' (2017) 'Cluster Mapping as a Tool for Development' demonstrates the structuring of territories in accordance with the clusters that are located on them and reflects the idea of their visual display in the form of interactive graphics. In the article by Falcioğlu and Akgüngör (2008), the authors carry out a cluster analysis of regions using data from Turkey and testing it in accordance with the industrial production facilities located on its territory. In their work, Feser and Bergman (2000) justify the concept of grouping regions in accordance with the main industry clusters that appear at the state level. They also highlight key cluster patterns that may be inherent at the federal level.

The review of the above-mentioned works is expressed in a detailed understanding of how certain specific state industry clusters or industrial production mechanisms (as the main tools for creating a material product) function, which are implemented in the country under consideration. As a result, the approaches to the management of territories used in the reviewed works constitute an empirical approximation and are inherently unique, specific and difficult to adapt for other spatial subjects.

In terms of technical analysis, we use different variations of correlation analysis to determine whether there is a relationship between socio-economic metrics. The limitation of their application for most economic samples lies in the lack of data uniformity. Consequently, relationships, as such,

cannot be unambiguously detected but, with the appropriate grouping of objects included in the sample, it is possible to model stable patterns within each group.

The use of the perceptron model on the display area of the quantitative data, which, by its very nature, allows us to linearly divide the n -dimensional space into two components in accordance with the manifestation of the concentration of statistical estimates of interest in them, can represent a possible solution to this problem. A feature of this approach is that, unlike the classical versions of cluster analysis, it allows us to form groups by linearly dividing them and not around the point of accumulation of data, which in turn allows for a more correct display of the dynamics of the process.

The implementation of managed territorial object clustering is reflected in the implementation of ‘sustainable economic development’ concepts. The fundamental work of Uskova (2009), the ‘Management of Sustainable Development of the Region’ monograph, touches on this topic. In it, she considers these things through the prism of Russian regions and their smaller structural units—municipalities. Another article, written by a team of authors under the leadership of Pozdnyakova (Pozdnyakova et al., 2017), also demonstrates the importance of the proper clustering of territories for the formation of stable signs of development and for the growth of economic processes and phenomena within them. There is also an emphasis on the fact that the grouping of territories should be based on innovations, the importance of which we mentioned earlier in the ‘Introduction’ section of this article. Furthermore, a scientific work by Rentkova (2019) shows the importance of proper clustering of territorial objects (the manuscript focuses on cities, using the example of the Republic of Slovakia) in implementing the territories’ principles of sustainable economic development.

3. Materials and methods

The basic functional unit of artificial neural networks (ANNs) is a formation such as a perceptron (a single-layer artificial neural network) (Shamin, 2019, n.d.). Its discovery occurred around 1950s and is associated with Rosenblatt (1962), where a principal point that should be noted is its ‘learning’ property, which seemed to be very promising at first. Subsequently, Minsky and Papert (1969) showed the limitations of this object (some of the simplest logical problems cannot be solved with it) in their works, which led to a decline of interest in this tool. Its schematic illustration is shown in Figure 1:

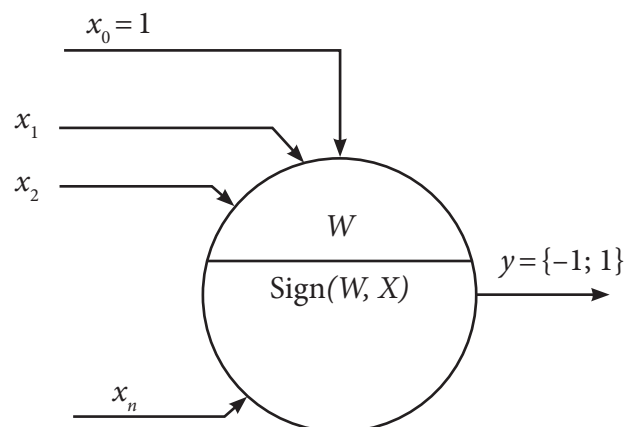


Figure 1. Perceptron circuit (compiled by the authors)

Here, $i = \overline{0, n}$; $n \in \mathbb{N}$ is the set of inputs to the perceptron body; $x_i \parallel X = \{x_0, x_1, \dots, x_n\}$ is the value supplied to the i -th input; $x_0 = -1 \parallel 1$ is the dummy input, the value of which is -1 or 1 ; $x_i \in \mathbb{R}$ represents user inputs, the estimates of which can take values from a set of real (real) numbers; $w_i \parallel W = \{w_0, w_1, \dots, w_n\}$ – weight coefficients;

$$\text{Sign}(t) = \begin{cases} -1, & t < 0, \\ 1, & t \geq 0. \end{cases}$$

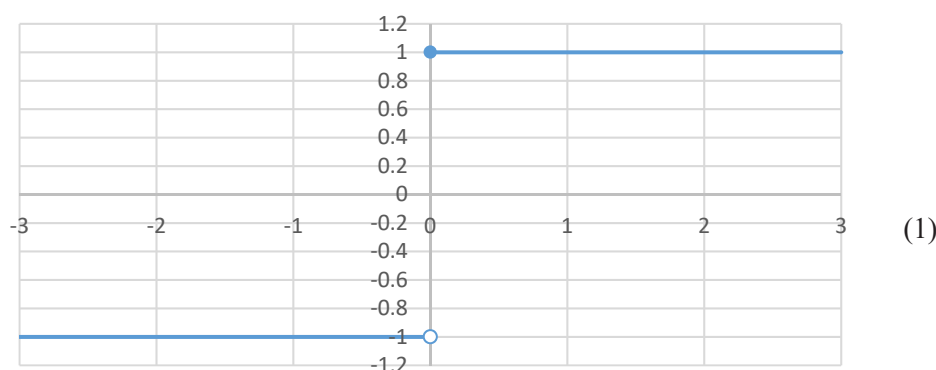


Figure 2. Heaviside step function (compiled by the authors)

Here, $\text{Sign}(t)$ is the activation function of iteration t ; $y = \text{Sign}(W, X)$ is the output value of the perceptron, resulting from calculating the Heaviside step function (Figure 1) from the inner product; $(W, X) = w_0x_0 + w_1x_1 + \dots + w_nx_n = \sum_{i=0}^n w_i x_i$ is the inner product of \overline{W} and \overline{X} . For $x_0 = 1$ (Figure 1), the inner product will take the following form – $(W, X) = w_0 + w_1x_1 + \dots + w_nx_n = w_0 + \sum_{i=1}^n w_i x_i$.

In this case $y \in \{-1; 1\}$, i.e. the perceptron performs binary classification between vectors. If we do not want the classification to be binary, then Sign does not apply. In this case, we do not determine the class but with what force the considered value belongs to a particular class.

The key thing about the perceptron is that the values of vector W can change as we work with it. This process is called learning in the discipline, i.e. we adjust the values of vector W in the way that we need (in accordance with the original data).

Learning, in turn, is divided into two main directions: 1) supervised learning (the training set is labelled, i.e. the correct answer is given to the and 2) unsupervised learning. Supervised learning is a typical task statement for ANN. The initial data for it is presented in the following table (Table 1).

Table 1 A priori data set for training a perceptron on labelled data (supervised learning), where m is the number of observations in the set (compiled by the authors)

x_1^1	x_2^1	...	x_n^1	y^1
x_1^2	x_2^2	...	x_n^2	y^2
...
x_1^m	x_2^m	...	x_n^m	y^m

3.1 Detailed perceptron learning algorithm

First, let us set the initial values for vector W . For example, $W = \vec{0}$. The values can also be selected at random. This affects the rate of convergence of the perceptron, provided it is present. Second, we repeat the procedure described below many times (the number of repetitions is selected experimentally):

1) In accordance with $j = \overline{1, m}$; (j is a certain number of our observation), we calculate d :

$$d = \text{Sign}(W, X) = \begin{cases} -1, & \text{if } (W, X) < 0, \\ 1, & \text{if } (W, X) \geq 0. \end{cases}$$

How j is selected is an open question. There is an option to select it sequentially (if it was previously distributed in an arbitrary order) or stochastically. In accordance with the practice of its own implementation, the random sorting of objects that are divisible by the perceptron should be laid down in the form of a certain iteration. In this case, we randomly sort the trained set until it gives a certain specified result (e.g. splitting the population under study into an acceptable percentage).

2) If $d \cdot y = -1$, then the recognition is performed incorrectly and it is necessary to adjust the values of W :

$$w_i = w_i + \alpha \cdot y \cdot x_i, \quad \alpha > 0.$$

α is a parameter that sets the rate of our learning, and is determined experimentally. Traditionally, it is positive and small. The smaller it is, the more accurately we learn, but longer and vice versa. If $d \cdot y$ is still -1 , then we continue to adjust the weights until we obtain the correct answer. We proceed to the next observation and repeat what we did in steps 1 and 2. The calculation according to the described algorithm is presented below (Table 2).

Thus, the perceptron model under the given conditions will have the following form:

$$\text{Sign}(0.15 - 3.7443x_1).$$

Table 2 Algorithm for calculating weights for the perceptron model y from x (compiled by the authors), where $X = \{x_0, x_1\}$; $x_0 = \{1, 1, 1, 1\}$; $x_1 = \{0.6622, 74.998, 8.9736, 0.0281\}$; $y = \{1, -1, -1, 1\}$; $\alpha = 0.05$

x_0	x_1	y	w_0	w_1	(W, X)	d	$d \cdot y$
1	0.6622	1	0	0	$=w_0 \cdot x_0 + w_1 \cdot x_1 = 0 \cdot 1 + 0.06622 = 0$	1	1
1	74.9981	-1	0	0	0	1	-1
1	74.9981	-1	$=w_0 + \alpha \cdot y \cdot x_0 = w_0 + \alpha \cdot y = 0 + 0.05 \cdot (-1) = -0.06$	$=w_0 + \alpha \cdot y \cdot x_1 = 0 + 0.05 \cdot (-1) \cdot 74.9981 = -3.7499$	-281.2860	-1	1
1	8.9736	-1	-0.0500	-3.7499	-33.7000	-1	1
1	0.0281	1	-0.0500	-3.7499	-0.1553	-1	-1
1	0.0281	1	0.0000	-3.7485	-0.1053	-1	-1
1	0.0281	1	0.0500	-3.7471	-0.0553	-1	-1
1	0.0281	1	0.1000	-3.7457	-0.0052	-1	-1
1	0.0281	1	0.1500	-3.7443	0.0448	1	1

In conclusion, we want to note two main properties of the perceptron: 1) linear division of the set into two classes and 2) generalisation, expressed in the fact that despite the possibility of incorrect data, its work will be reliable in general. It is also worth noting that the success of using an artificial neural network is ensured by a good learning set. The solution can be the generation of tests (e.g. in a branch of knowledge such as mechanics).

The above algorithm is largely iterative. It can set specific parameters that, due to the simplicity of the pilot simulation, are indicated in the form of constants, taken in accordance with the recommendations of leading scientists in this field. They can also be made dynamic as part of the further development of the study or the initial values can be set in accordance with the actual, current conditions of the problem under consideration.

In the case of the perceptron operation algorithm, the following rules can be set:

- we denoted the vector of weight W as $\vec{0}$. However, if you choose any specific value, then the location of the hyperplane that divides the hyperspace into two parts will be closer to the desired one and, therefore, the learning process will be faster;

- in our case, the learning rate parameter α is taken at the level of 0.05, as a kind of positive practice in applied research related to the perceptron. At the same time, it is constant. However, it is still possible to make it dynamic and to either speed up or slow down the process of finding the acceptable weight. It can also be set separately for each variable included in the modelled structure;

- in accordance with the data included in the training set, the final model of the generated perceptron may be slightly different and divide the studied population without generalisation. In this regard, it is important to set more stringent modelling requirements or to carry out a procedure for mixing observations until the final result meets the specified conditions.

The things mentioned above are the ones primarily considered in two fundamental works: ‘Principles of Neurodynamic’ (Rosenblatt, 1962) and ‘Perceptrons’ (Minsky and Papert, 1969). In the case of working out any complex specific nuances of these algorithms, their use should be carried out manually, modelling each of the possible aspects independently in a computer environment. However, in the case of reproducing experiments that have already been tested or are largely similar to them in terms of the conceptual part, ready-made tools are also suitable – for example, various Python libraries, such as Keras or TensorFlow. An even more narrowly focused option is the neural network toolkit of the Statistica software, maintained by Stata software.⁵

3.2 Perceptron learning algorithm using Python tools

The implementation listing of the perceptron identified above, which divides the labelled training set into two classes (‘1’ or ‘-1’), is provided below (Figure 2).

The parameters w_0 , w_1 and α are set by the researcher independently and can be selected under the conditions of the problem. The metric α for each weight can be unique and, for better convergence, is set in terms of acceleration rather than constant rate.

If the data under study is not previously labelled, then the implementation of the perceptron may look like this (Figure 3):

As in the first listing (Figure 2), the parameters w_0 , w_1 and α can be set in accordance with the specifics of the data under study. In addition to this, you can apply further normalisation of quantitative estimates to reduce the impact of the response of numerical values supplied to the input of the algorithm during training on the modelling of weight coefficients.

⁵ Stata: Software for Statistics and Data Science. <https://www.stata.com/>

```

1  for i in range(len(x1)):
2      d = -1
3
4      while d == -1:
5          b = w0*x0[i] + w1*x1[i]
6          if b < 0:
7              b = -1
8          else: b = 1
9          d = b*y[i]
10         if d == -1:
11             w0 = w0 + a*y[i]*x0[i]
12             w1 = w1 + a*y[i]*x1[i]
13         else:
14             w0 = w0*1
15             w1 = w1*1

```

Figure 2. Algorithm for the implementation of the perceptron on labelled data (output value ‘1’ or ‘-1’) (compiled by the authors)

Note: Parameters: w_0 , w_1 are the weights of the variables; a is the parameter responsible for the rate of change of the simulated weights W . Variables: x_0 , x_1 are the vectors of values supplied to the input.

```

1  for i in range(len(x1)):
2      y_prog = w0*x0[i] + w1*x1[i]
3
4      if y_prog > y[i]:
5          while y_prog > y[i]:
6              w0 = w0 - a*x0[i]
7              w1 = w1 - a*x1[i]
8              y_prog = w0*x0[i] + w1*x1[i]
9
10         elif y_prog < y[i]:
11             while y_prog < y[i]:
12                 w0 = w0 + a*x0[i]
13                 w1 = w1 + a*x1[i]
14                 y_prog = w0*x0[i] + w1*x1[i]
15
16         else:
17             w0 = w0*1
18             w1 = w1*1

```

Figure 3. Algorithm for the implementation of the perceptron (compiled by the authors)

Note: Parameters: w_0 , w_1 are the weights of the variables; a is the parameter responsible for the rate of change of the simulated weights W . Variables: x_0 , x_1 are the vectors of values supplied to the input.

3.3 Perceptron clustering algorithm on unlabelled data

1. Based on the available data, we construct a model of paired linear regression (Seber, 1977) and calculate its coefficient of determination. In accordance with its ratio and the levels of the Chaddock scale (Koterov et al., 2019, p. 14), we set an acceptable level of model accuracy for us. For example, 0.7 for the Pearson correlation (in the work, when tested on empirical data, the critical level is set at 0.8), described in one of the scientific papers referring to Chaddock as characterising a ‘very good relationship’. At this level, the variance of one variable in relation to the other begins to exceed 50%. If this condition is satisfied, no clustering is required. If not, then go to step 2.
2. We sort the training sample randomly.
3. We train the perceptron according to the scheme shown in the listing figures (Figure 3).
4. In accordance with the obtained linear clustering model, we divide the sample population into two parts. In this case, the ratio of the two new aggregates must meet the following specified criteria:
 - 1) The number of observations in one of the newly formed populations must be greater than or equal to the specified size of the original population (in our example, we set this parameter at the level of 20%);

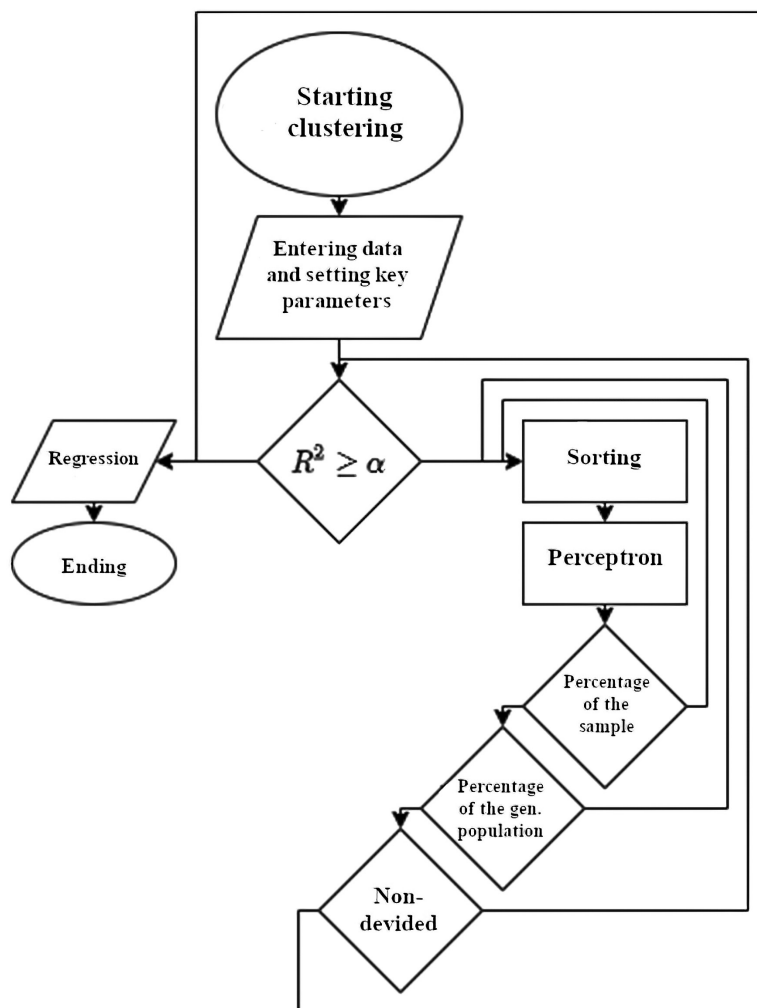


Figure 4. Clustering algorithm (compiled by the authors)

- 2) The number of observations in one of the newly formed populations must also be simultaneously greater than or equal to the specified size of the general population (in our example, we set this parameter at the level of 5%).

In case of non-compliance with one of the two above-mentioned criteria, we return to step 2. If the conditions are satisfied, we move on.

5. We check the newly formed groups for the possibility of further division in accordance with requirement 2) indicated in step 4. To do this, each of these groups must be divided in half. If the result from the division does not satisfy 2), then the clustering for the original group is completed and the final model of paired linear regression can be built on it through analogy with the one indicated at the first step of the algorithm. If the newly formed group can be divided, then check it for the condition R^2 . If the condition is satisfied, no further

Table 3 The ratio of the average monthly salary to the cost of R&D per 10 thousand people, 2015–2019 (comparable prices according to the consumer price index)

Code	Region	R&D costs per 10 thousand people, million rubles	Average monthly salary, rubles
1	Belgorod region (2015)	14.38	29.544
2	Bryansk region (2015)	5.19	25.161
...
400	Sakhalin region (2019)	21.34	84.872

Note: Compiled by Regions of Russia. Socio-economic indicators. 2020: P32 Stat. sat., Moscow, 2020. <https://rosstat.gov.ru/folder/210/document/13204>

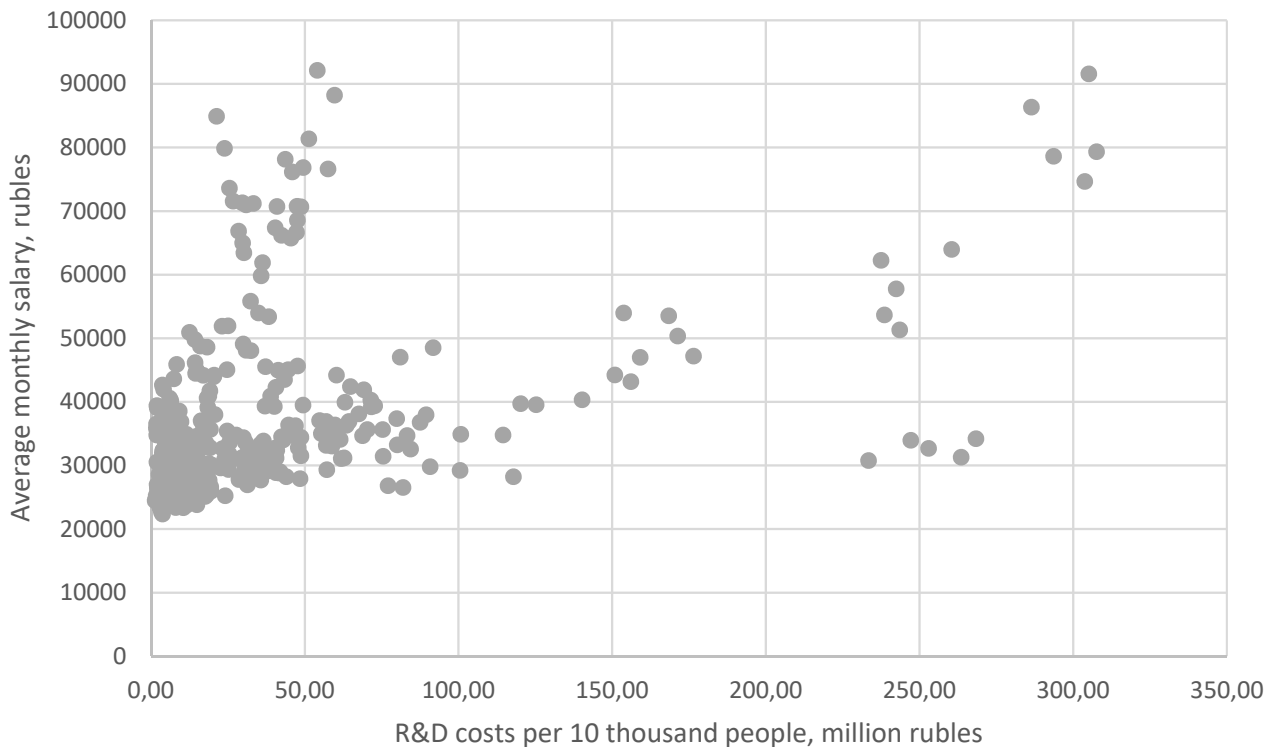


Figure 5. The ratio of the average monthly salary to the cost of R&D per 10 thousand people, 2015–2019 (comparable prices according to the consumer price index)

Note: Compiled by Regions of Russia. Socio-economic indicators. 2020: P32 Stat. sat., Moscow, 2020. <https://rosstat.gov.ru/folder/210/document/13204>

clustering is required. For this group, we build a model of paired linear regression in accordance with the one indicated in the first step of the algorithm (in fact, it is a return to step 1). If the condition is not satisfied, then we skip the newly formed group in accordance with all the steps of the algorithm and so on, until we get groups that cannot be divided or until the data set that is contained in them does not correspond to the set determination coefficient.

For clarity, the developed scheme of the algorithm is presented in Figure 4 below.

The presented algorithm is a generalisation of the numerical methods indicated before it. It can be detailed in the 'Data entry' part and the 'Regression' part. The initial data for testing the methods indicated in the work are presented in Table 3 below and are fully reflected in Figure 5.

Similarly, as in Table 3 and Figure 5, we make comparisons for 'R&D costs per 10 thousand people of the population' with 'average per capita income per month', 'GRP' and 'innovation activity'. We bring monetary indicators to a single point of reference in time through the consumer price index.

4. Results

Using perceptron clustering, we construct paired linear regression models, showing the linear response of investments and expenditures on science to one of the four indicators identified in the work for each of the groups formed. The visualisation of the performed calculations is presented below (Tables 4–7 and Figures 6–9).

Table 4. Detailed clustering procedure using the example of the statistical dependence of the average monthly salary on R&D costs (*compiled by the authors*)

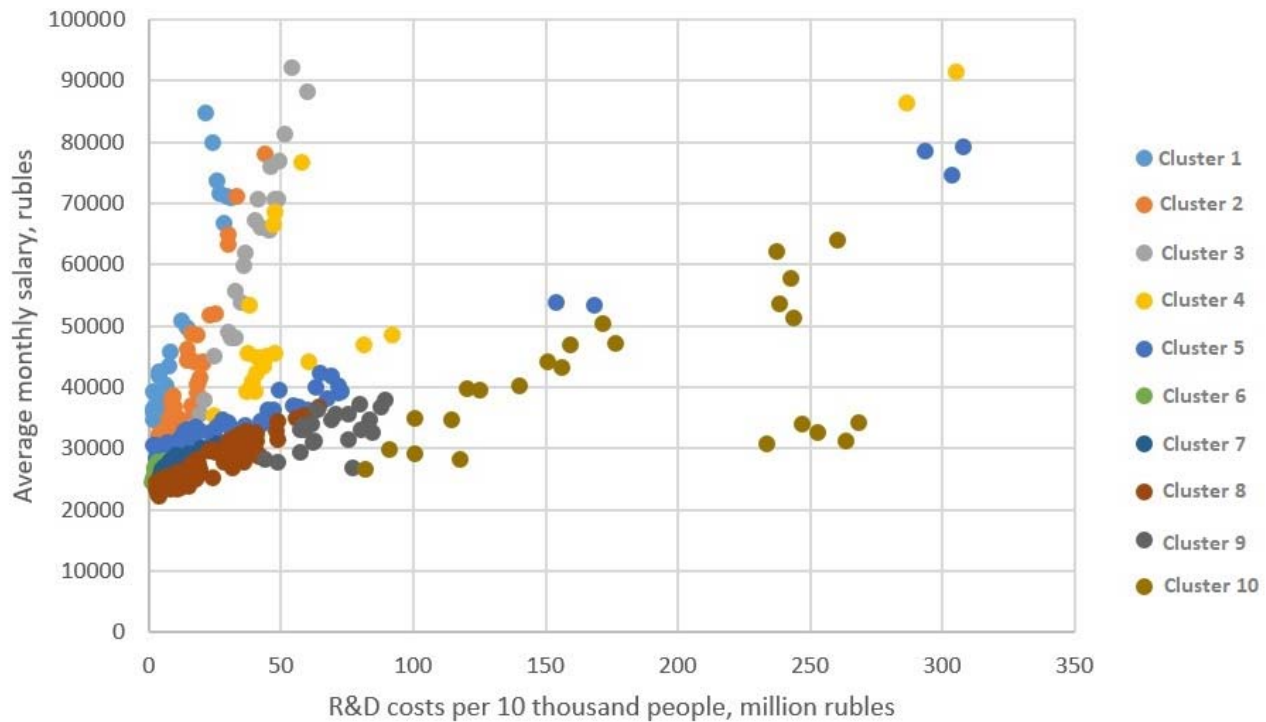
First iteration (one cluster)		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
400	$y=32\,177.8834+109.8512x$	0.1996
<i>Perceptron: $y=27\,348.05+149.36x$</i>		
Second iteration (two clusters)		
1.1. First cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
200	$y=37\,238.5282+173.9474x$	0.3107
<i>Perceptron: $y=30\,267.15+193.607x$</i>		
1.2. Second cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
200	$y=26\,047.2989+87.7247x$	0.6436
<i>Perceptron: $y=23\,988.9+246.5195x$</i>		
Third iteration (four clusters)		
1.1.1. First cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
122	$y=26\,047.2989+87.7247x$	0.5911
<i>Perceptron: $y=41\,336.8574+236.7216x$</i>		
1.1.2. Second cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
78	$y=28\,701.2019+160.4252x$	0.9836

Table 4 (continued)

1.2.1. Third cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
59	$y=25\,477.1075+207.1883x$	0.7051
<i>Perceptron: $y=20\,642+1227.2615x$</i>		
1.2.2. Fourth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
141	$y=25\,860.0886+88.8163x$	0.6175
<i>Perceptron: $y=11\,826.8+387.581x$</i>		
Fourth iteration (seven clusters)		
1.1.1.1. First cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
30	$y=31\,932.627+1506.5539x$	0.8656
1.1.1.2. Third cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
92	$y=40\,291.616+231.0619x$	0.3834
<i>Perceptron: $y=27\,654.25+1\,310.948x$</i>		
1.2.1.1. Fourth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
26	$y=24\,352.1053+655.074x$	0.4881
1.2.1.2. Fifth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
33	$y=25\,089.2624+230.0391x$	0.8253
1.2.2.1. Sixth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
96	$y=23\,059.238+206.0468x$	0.8536
1.2.2.2. Seventh cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
45	$y=27\,790.1666+74.8948x$	0.3573
<i>Perceptron: $y=11\,826.8+387.581x$</i>		
Fifth iteration (nine clusters)		
1.1.1.2.1. Third cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
50	$y=23\,930.1484+1176.487x$	0.8633
1.1.1.2.2. Seventh cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
42	$y=27\,790.1666+74.8948x$	0.2525
<i>Perceptron: $y=9\,197.8+1\,575.455x$</i>		
1.2.2.2.1. Eighth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
21	$y=23\,590.6379+137.5564x$	0.3128
1.2.2.2.2. Ninth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
24	$y=27\,576.31+75.5219x$	0.1987
Sixth iteration (ten clusters)		
1.1.1.2.2.1. Seventh cluster		

Table 4 (finished)

<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
22	$y=9\ 187.076+1\ 370.7199x$	0.948
1.1.1.2.2.2. Tenth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
20	$y=40\ 194.7432+165.6774x$	0.6153

**Figure 6** Clustering payroll with R&D costs (*compiled by the authors*)

The main characteristics of the clusters formed with the ratio of wages and R&D costs, as well as their graphical visualisation, are presented in Table 4 and Figure 6 above. With less detail, the results of calculations regarding the relationship between R&D costs and per capita income, GRP and innovation activity are presented below (Tables 5–7 and Figures 7–9).

Table 5. The result of clustering on the example of the statistical dependence of average per capita income on R&D costs (*compiled by the authors*)

First iteration (one cluster)		
1. <i>Perceptron</i> : $y=22\ 014.6+391.243x$		
Second iteration (two clusters)		
1.1. <i>Perceptron</i> : $y=19\ 570.6+636.3575x$		
1.2. <i>Perceptron</i> : $y=17\ 694.2+574.5665x$		
Third iteration (four clusters)		
1.1.1. First cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
123	$y=23\ 965.7267+665.5832x$	0.8352
1.1.2. Second cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>

Table 5 (continued)

24	$y = -29.7756 + 0,0016x$	0.24
1.2.1. Perceptron: $y = 13\,248.8 + 1492.3185x$		
1.2.2. Perceptron: $y = 14\,230.65 + 188.6715x$		
Fourth iteration (six clusters)		
1.2.1.1. Third cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
22	$y = 20\,174.1019 + 1483.178x$	0.4043
1.2.1.2. Fourth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
22	$y = 20\,028.4187 + 470.0996x$	0.7823
1.2.2.1. Perceptron: $y = 14\,323.85 + 769.799x$		
1.2.2.2. Fifth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
35	$y = 15\,730.7867 + 120.8396x$	0.5997
Fifth iteration (seven clusters)		
1.2.2.1.1. Sixth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
21	$y = 15\,869.667 + 715.4577x$	0.956
1.2.2.1.2. Perceptron: $y = 1\,567.8 + 776.533x$		
Sixth iteration (eight clusters)		
1.2.2.1.2.1. Seventh cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
83	$y = 14\,312.6826 + 505.4476x$	0.805
1.2.2.1.2.2. Eighth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
70	$y = 21\,042.7696 + 164.9877x$	0.8656

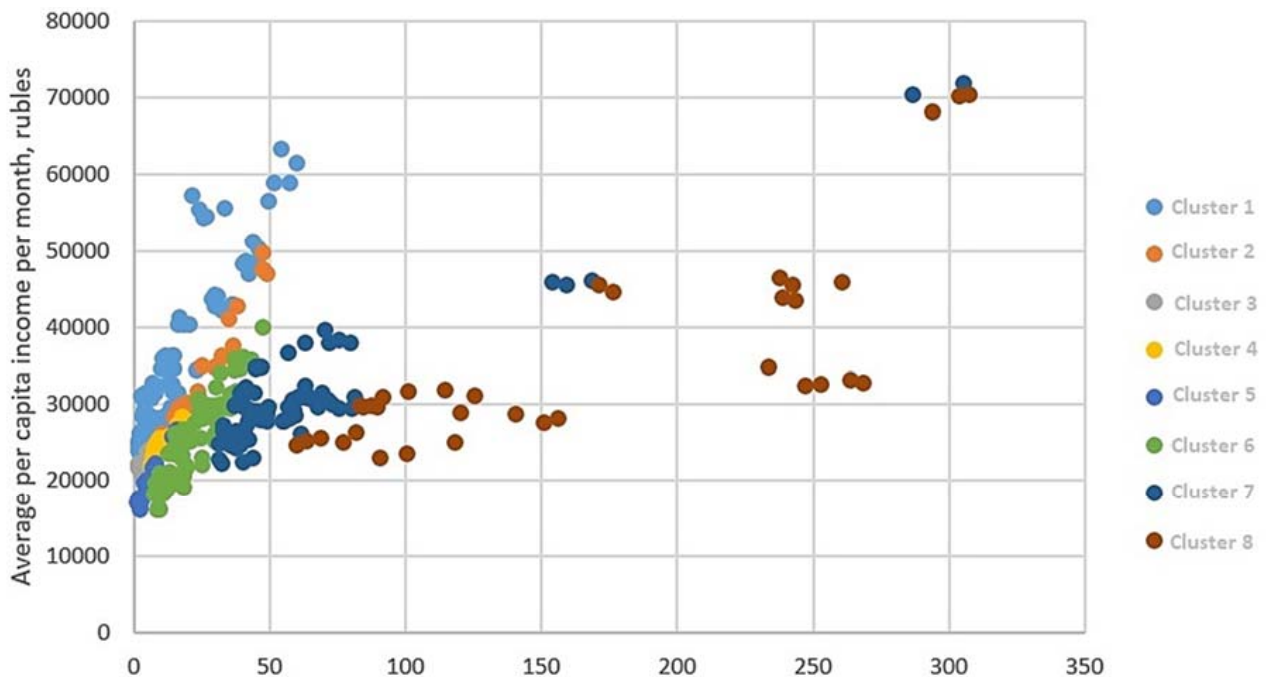


Figure 7 Clustering of average per capita incomes with R&D costs (compiled by the authors)

Table 5 and Figure 7 show the results of the modelling cluster analysis of regions with the ratio of their average per capita income and R&D costs. The performed calculations can be considered successful because most of the obtained models of the growth of average per capita income on R&D expenditure dependence have a high coefficient of determination ($R^2 \geq 0.8$).

Table 6. The result of clustering on the example of the statistical dependence of GRP per 10 thousand people population on R&D costs (*compiled by the authors*)

First iteration (one cluster)		
1. Perceptron: $y = 2\,864.2 + 18.704x$		
Second iteration (two clusters)		
1.1. Perceptron: $y = 3\,238.2 + 337.842x$		
1.2. Perceptron: $y = 1\,962.95 + 30.5355x$		
Third iteration (four clusters)		
1.1.1. First cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
17	$y = 4\,305.258 + 401.6492x$	0.833
1.1.2. Perceptron: $y = 2\,102.4 + 156.7485x$		
1.2.1. Perceptron: $y = 1\,472.15 + 95.059x$		
1.2.2. Perceptron: $y = 936.15 + 44.042x$		
Fourth iteration (seven clusters)		
1.1.2.1. Second cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
51	$y = 3\,067.3004 + 175.3168x$	0.8594
1.1.2.2. Perceptron: $y = 171.3 + 74.482x$		
1.2.1.1. Perceptron: $y = 1\,002.1 + 257.4945x$		
1.2.1.2. Third cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
24	$y = 2\,204.2446 + 29.0543x$	0.8895
1.2.2.1. Perceptron: $y = 837.25 + 138.5545x$		
1.2.2.2. Fourth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
23	$y = 2\,810.7234 + 9.4134x$	0.2974
Fifth iteration (ten clusters)		
1.1.2.2.1. Perceptron: $y = 136.1 + 223.9675x$		
1.1.2.2.2. Fifth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
21	$y = 2\,234.1708 + 31.7977x$	0.8511
1.2.1.1.1. Sixth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
28	$y = 2\,846.8172 - 92.2809x$	0.1194
1.2.1.1.2. Seventh cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
24	$y = 1\,939.6131 + 72.9646x$	0.8478
1.2.2.1.1. Eighth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
19	$y = 1\,030.987 + 137.5103x$	0.9065

Table 6 (continued)

1.2.2.1.2. Ninth cluster		
Number of observations	Regression	R^2
26	$y=1\,468.6517+38.3244x$	0.8162
Sixth iteration (eleven clusters)		
1.2.2.1.2.1. Tenth cluster		
Number of observations	Regression	R^2
17	$y=1\,825.8237+152.8044x$	0.8634
1.2.2.1.2.2. Perceptron: $y=23.4+153.024x$		
Seventh iteration (twelve clusters)		
1.2.2.1.2.2.1. Eleventh cluster		
Number of observations	Regression	R^2
22	$y=642.3635+156.6956x$	0.9308
1.2.2.1.2.2.2. Perceptron: $y=8.05+109.579x$		
Eight iteration (thirteen clusters)		
1.2.2.1.2.2.2.1. Twelfth cluster		
Number of observations	Regression	R^2
27	$y=707.2714+109.6994x$	0.7761
1.2.2.1.2.2.2.2. Thirteenth cluster		
Number of observations	Regression	R^2
21	$y=520.8359+85.6223x$	0.831

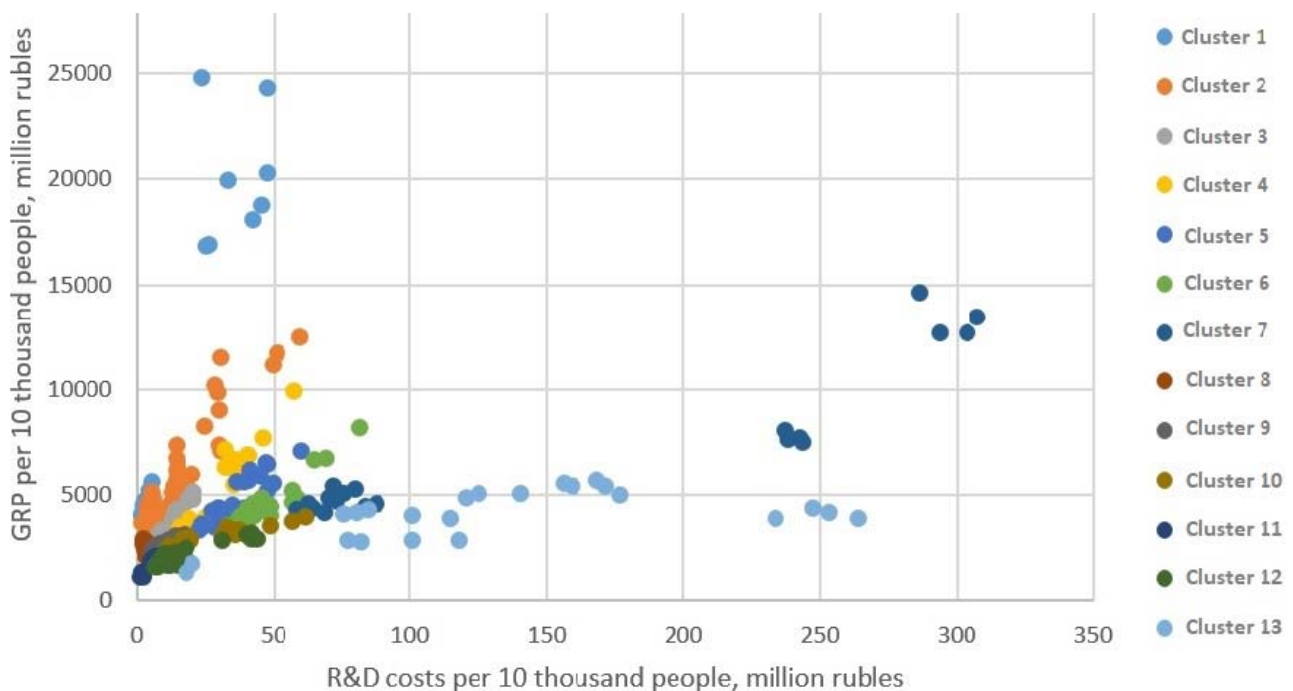


Figure 8. GRP clustering with R&D costs (compiled by the authors)

Table 6 and Figure 8 present the results of modelling the clustering of regions in determining the relationship between the simultaneous growth of R&D costs and GRP. As in the previous versions, the algorithm showed a good result, simulating most of the dependencies at the $R^2 \geq 0.8$ level.

Table 7. The result of clustering on the example of statistical innovation activity on R&D costs
(compiled by the authors)

First iteration (one cluster)		
1. Perceptron: $y=4.8+0.1818x$		
Second iteration (two clusters)		
1.1. Perceptron: $y=5.9+0.1864x$		
1.2. Perceptron: $y=2.4+0.2573x$		
Third iteration (four clusters)		
1.1.1. Perceptron: $y=5.295+0.3286x$		
1.1.2. First cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
41	$y=5.3453+0.181x$	0.9849
1.2.1. Second cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
41	$y=3.4789+0.2323x$	0.8712
1.2.2. Perceptron: $y=1.39+0,156x$		
Fourth iteration (six clusters)		
1.1.1.1. Perceptron: $y=4.925+1.1323x$		
1.1.1.2. Third cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
25	$y=6.4438+0.2067x$	0.8804
1.2.2.1. Fourth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
102	$y=2.2743+0.1801x$	0.9046
1.2.2.2. Perceptron: $y=1.39+0,156x$		
Fifth iteration (eight clusters)		
1.1.1.1.1. Fifth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
20	$y=8.4344+1.1673x$	0.6182
1.1.1.1.2. Perceptron: $y=1.615+0.9766x$		
1.2.2.2.1. Perceptron: $y=0.315+0.1352x$		
1.2.2.2.2. Sixth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
21	$y=-1.4109+0.0685x$	0.3228
Sixth iteration (ten clusters)		
1.1.1.1.2.1. Seventh cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
34	$y=3.3653+0.979x$	0.8472
1.1.1.1.2.2. Eighth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
30	$y=6.6578+0.3444x$	0.892
1.2.2.2.1.1. Ninth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
25	$y=0.7996+0.1498x$	0.9717
1.2.2.2.1.2. Tenth cluster		
<i>Number of observations</i>	<i>Regression</i>	<i>R²</i>
51	$y=0.4622+0.0991x$	0.8776

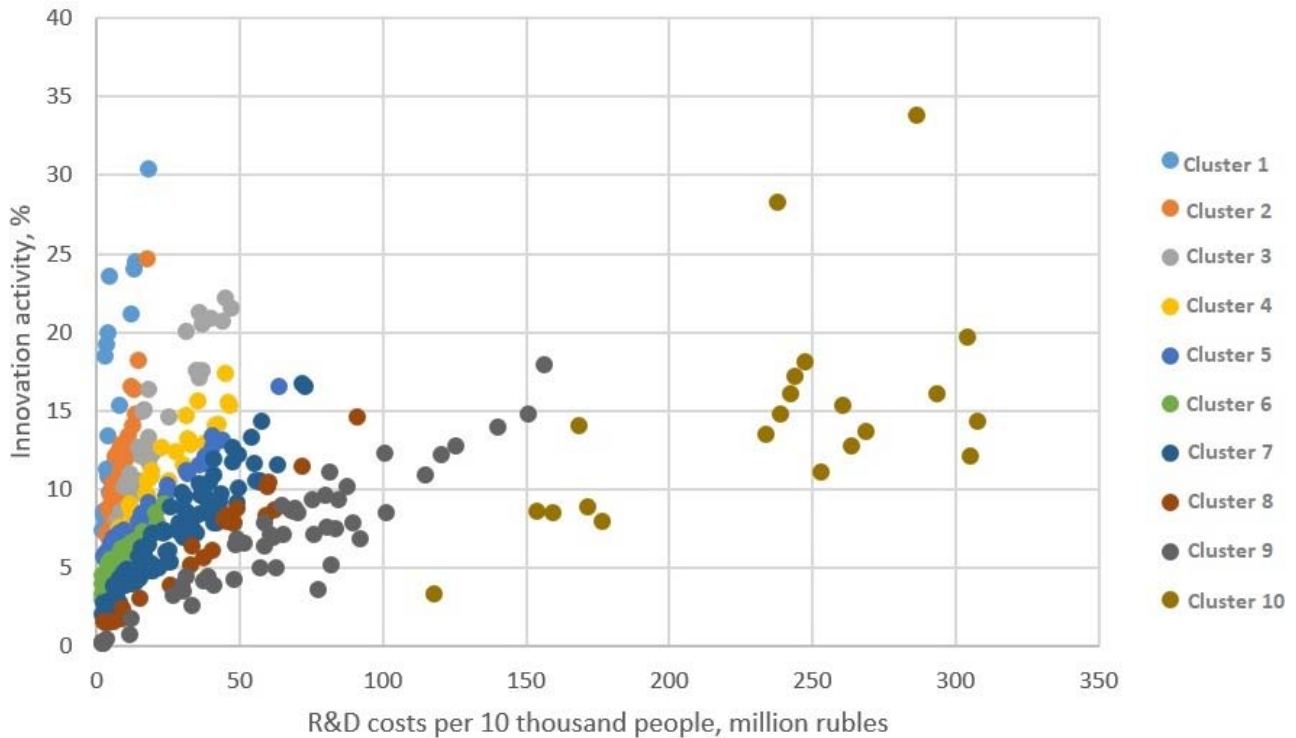


Figure 9. Clustering innovation activity with R&D costs (*compiled by the authors*)

In the last clustering, the learning rate of the weights was reduced by one order of magnitude ($\alpha=0.05 \rightarrow \alpha=0.005$). The need for this procedure arises from the fact that the perceptron could not divide the population supplied to it as input into two parts in accordance with the condition of a sufficient share of the sample and the general population. This is due to the size of the indicators involved in the learning for which the weights are modelled on the resulting response. For wages, per capita income and GRP per unit of population, the average dimension is measured in thousands of units, for innovation activity is measured in dozens. A possible universal way to implement the perceptron algorithm is to pre-normalise the data.

5. Discussion

Forecasting estimates for socio-economic systems is a complex and urgent task in view of the disparate behaviour of the relationships between them in the field of their representation. In contrast to natural systems, socio-economic patterns visually often have several variants of development. To some extent, this may be due to the fact that the objects of research that are identical for us are actually not identical. A variant of this can be territorial entities that are nominally designated as regions (municipalities, states, countries and other similar objects can also appear here), although, in fact, they are something different.

It is also worth noting here that socio-economic information is often unstable, even for identical objects, in contrast to natural science data. If we measure the mass of a body or, for example, its mechanical speed of movement, then we can compare it with another object using these same characteristics. In the case of economy, things are more complicated. Not only does the measurement of certain socio-economic characteristics largely depend on the opinion of the person who

takes these indicators but the indicators themselves are to some extent dynamic in nature. An example of this is a currency that, when used in a different areas, will have different purchasing power. A possible option for more accurate modelling of such processes can be quantum computing (Kozyrev, 2018). In one of the most recent publications in 2020, a team of authors led by Moreira (Moreira et al., 2020) proposed a universal scheme for modelling the decision-making process that allows us to reflect the irrationality of human behaviour and thinking. The complexity of modelling socio-economic processes and the inefficient use of existing mathematical methods in relation to them is shown in the work of Martínez-Martínez (2014). An alternative solution to them is quantum computing.

The use of the perceptron model in this work allowed us to divide the studied population in a universal manner in accordance with the behaviour of the dynamics of three different indicators of the socio-economic well-being of citizens in response to changes in the R&D cost amount. In general, the trend in all four metrics (salary, per capita income, GRP and innovation activity) with an increase in the amount of spending on science can be described as positive, however, it manifests itself differently in different regions. For some it is faster, for some it is slower.

The final linear regression models have a high coefficient R^2 , greater than 0.8, which, in accordance with established econometric practice, is a good result that can be used in applied management activities. At the same time, in the future, the model proposed in this paper will have the potential for improvement in the form of connecting a variation of the genetic algorithm to it when choosing the best possible clustering option. The linear regression model can also be replaced with a function that more closely approximates the actual data: exponential trend, if there is an acceleration of the dynamics of the process under study; logarithmic, if there is a damping; trigonometric, if there are static fluctuations.

Modelling the impact of investments in science is an important component for planning the qualitative development of human society because science constitutes the 'spark of ignition' when creating new technologies or innovations. The forecast of the response and return from it would allow us to invest into various branches of knowledge with the greatest efficiency in order to obtain the best result at the end. In addition, it becomes possible to take a more selective approach to the management of individual territories in the entire totality of the controlled system in order to implement socially significant economic effects in a manner that is sustainable for them in the long term.

6. Conclusion

In accordance with the set goal, it can be concluded that the algorithm proposed in the study, based on the perceptron model, allows us to successfully cluster territorial objects for purposes of further modelling of correct dependencies of the socio-economic metrics found in them. Amongst the positive features of the proposed algorithm, it is worth noting its universality.

Furthermore, in this study, we obtained the following results:

The results of earlier research in the direction of clustering of territorial objects were generalised and systematised. This allowed us to identify aspects such as: 1) the lack of universal cluster analysis methods for territories and the fact that their grouping is based on the specifics of industry clusters located on them and large industrial facilities; 2) the main tools used in such studies constitute different variations of correlation analysis, which does not give unambiguous answers with different types of information being studied.

The clustering algorithm based on the application of the perceptron model allowed us to divide the data set under study in such a way that we could model monotonically increasing or decreasing dependencies inside them.

The use of the developed algorithm successfully proved itself when tested on Rosstat statistical data on investments in R&D, wages and average per capita income, GRP and innovation activity. The experiment provided a good result, confirmed by the majority of finite linear regression models with a determination coefficient of 0.8 units and higher.

The models constructed in the work can be used within specific territories of Russia, allowing for the adjustment of the growth of wages and average per capita income of the population, GRP and innovation activity of companies in accordance with the monetary investments in science in these regional subjects.

The universality of the algorithm can be successfully applied in the construction of other functional dependencies of socio-economic indicators and for administrative territories of other countries.

One further development of this study could focus on the technical side and be expressed in the refinement of the clustering algorithm via the introduction of a genetic algorithm and the building of more accurate final models based on the data included in the final clusters. Also, another development of the study could focus on the managerial side to determine the most favourable regions of the entire study population, represented by the territorial landscape of Russia, for purposes of scientific component development from which the best response to the growth of the well-being of the citizens living in these regions could be extracted.

The tools developed and used in this work can also be applied, using analogy, to other territorial entities.

References

- Alferyev, D.A., 2018. Forecasting of Indicators of the Scientific and Technological Space by Means of Spatial Econometric Models on the Example of the Republic of Belarus. in: *Belarusian Science in the Conditions of Modernization: Materials of the International Journal of Economics*. Scientific and Practical Conference, Minsk, September 20–21, 2018. StroyMediaProekt, Minsk, Belarus, pp. 115–118.
- Falcioğlu, P., Akgüngör, S., 2008. Regional Specialization and Industrial Concentration Patterns in the Turkish Manufacturing Industry: An Assessment for the 1980–2000 Period. *Eur. Plan. Stud.* 16, 303–323. <https://doi.org/10.1080/09654310701814678>
- Feser, E.J., Bergman, E.M., 2000. National Industry Cluster Templates: A Framework for Applied Regional Cluster Analysis. *Reg. Stud.* 34, 1–19. <https://doi.org/10.1080/00343400050005844>
- Fonotov, A., Bergal', O., 2020. Territorial Clusters in the System of Spacial Development: Foreign Experience. *Spat. Econ.* 16, 113–135. <https://doi.org/10.14530/se.2020.4.113-135>
- Ketels, C.H.M., 2017. *Cluster Mapping as a Tool for Development*. Institute for Strategy and Competitiveness-Harvard Business School: Boston, MA, USA, 52.
- Koterov, A., Ushenkova, L., Zubenkova, E., Kalinina, M., Biryukov, A., Lastochkina, E., Molodtsova, D., Vaynson, A., 2019. Strength of Association. Report 2. Graduations of Correlation Size. *Med. Radiol. Radiat. Saf.* 64, 12–24. <https://doi.org/10.12737/1024-6177-2019-64-6-12-24>
- Kozhevnikov, S.A., 2019. Spatial and Territorial Development of the European North: Trends and Priorities of Transformation. *Econ. Soc. Chang. Facts, Trends, Forecast* 6, 91–109. <https://doi.org/10.15838/esc.2019.6.66.5>
- Kozyrev, A.N., 2018. Quantum Economics and Quantum Computing in Economics. *Digital Economy* 3, 5–12.
- Kudryavtseva, T.Y., Skhvediani, A.E., 2020a. An econometric analysis of the regional industrial specialization: The Russian manufacturing industry case study. *Econ. Anal. Theory Pract.* 19, 1765–1790. <https://doi.org/10.24891/ea.19.9.1765>

- Kudryavtseva, T.Y., Skhvediani, A.E., 2020b. Studying Regional Clusters with the Use of Data Processing Systems: The Case of the Biopharmaceutical Cluster. *Regionology* 28, 48–79. <https://doi.org/10.15507/2413-1407.110.028.202001.048-079>
- Lugovaya, E.S., 2012. Innovation as the Basis for Modernizing Modern Society. *The Science Journal of Volgograd State University. Philosophy. Series 7, Sociology and Social Technologies* 2, 103–108.
- Martínez-Martínez, I., 2014. A connection between quantum decision theory and quantum games: The Hamiltonian of Strategic Interaction. *J. Math. Psychol.* 58, 33–44. <https://doi.org/10.1016/j.jmp.2013.12.004>
- Maslova, T.S., Lalaeva, A.A., 2018. Comparative Analysis of R&D Financing in Russia and Abroad. *Accounting in Budget and Non-Profit Organizations* 7, 2–10.
- Minakir, P.A., 2017. Theoretical Aspects of the Study of Spatial Economic Systems. *Journal of Economic Theory* 3, 7–10.
- Minsky, M., Papert, S., 1969. *Perceptrons: An Introduction to Computational Geometry*. M.I.T. Press, London.
- Moreira, C., Tiwari, P., Pandey, H.M., Bruza, P., Wichert, A., 2020. Quantum-like influence diagrams for decision-making. *Neural Networks* 132, 190–210. <https://doi.org/10.1016/j.neunet.2020.07.009>
- Pashintseva, N.I., 2018. Methodological Problems of Accounting and Statistics of Research, Development and Technological Works. *Questions of Statistics* 25, 66–72.
- Pozdnyakova, U.A., Popkova, E.G., Kuzlaeva, I.M., Lisova, O.M., Saveleva, N.A., 2017. Strategic Management of Clustering Policy During Provision of Sustainable Development, in: *Integration and Clustering for Sustainable Economic Growth*. Springer, Cham, pp. 413–421. https://doi.org/10.1007/978-3-319-45462-7_40
- Rentkova, K., 2019. The Clusters Phenomenon and Sustainable Regional Development. *IOP Conf. Ser. Mater. Sci. Eng.* 471. <https://doi.org/10.1088/1757-899X/471/10/102039>
- Rodina, V.V., 2014. Comparative Analysis of R&D Financing Mechanisms on the Example of Russia and the United States. *Monitoring Law Enforcement* 4, 65–73.
- Rosenblatt, F., 1962. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books, Washington, D.C.
- Schumpeter, J.A., 1980. *The Theory of Economic Development*. Routledge, London.
- Seber, G.A.F., 1977. *Linear regression analysis*. John Wiley and Sons, New York, N.Y.
- Seidl da Fonseca, R., Pinheiro-Veloso, A., 2018. The Practice and Future of Financing Science, Technology, and Innovation. *Foresight STI Gov.* 12, 6–22. <https://doi.org/10.17323/2500-2597.2018.2.6.22>
- Shamin, R.V., 2019. *Machine Learning in Economic Problems*. Green Print, Moscow.
- Shamin, R.V., n.d. Lecture 5: Training of the Perceptron in Motion Recognition Problems. October 8, 2019, 18: 30, Moscow, MIAN, room. 430 (Gubkina street 8)
- Stepanova, Yu.N., Lesnikova, M.S., 2017. The Role of Innovation in the Modern Development of Russian Society. *International Student Scientific Journal* 6.
- Uskova, T.V., 2009. *Managing the Sustainable Development of the Region: Monograph*. Institute of Socio-Economic Development of Territories of RAS, Vologda.
- Yegorenko, S.N., Bondarenko, K.A., Solovyova, S.V., 2018. Innovations: International Comparisons, in: *Human Development Report in the Russian Federation for 2018*. Analytical Center under the Government of the Russian Federation, Moscow, pp. 100–123.
- Yurchenko, N.Y., 2013. Financing of Research and Development Activities in Russia and Abroad. *Journal 'Humanitarian Bulletin' of BMSTU* 1(3), 1–11. <https://doi.org/10.18698/2306-8477-2013-1-33>
- Zhukovskaya, I.F., Ivlieva, N.N., Trufanova, S.A., 2021. Analysis of Direct and Indirect Methods of Stimulating Research and Development in the Russian Federation. *Problems of Management Theory and Practice* 1, 129–147.

СПИСОК ИСТОЧНИКОВ

- Falcioğlu, P., Akgüngör, S., 2008. Regional Specialization and Industrial Concentration Patterns in the Turkish Manufacturing Industry: An Assessment for the 1980–2000 Period. *Eur. Plan. Stud.* 16, 303–323. <https://doi.org/10.1080/09654310701814678>
- Feser, E.J., Bergman, E.M., 2000. National Industry Cluster Templates: A Framework for Applied Regional Cluster Analysis. *Reg. Stud.* 34, 1–19. <https://doi.org/10.1080/00343400050005844>
- Fonotov, A., Bergal', O., 2020. Territorial Clusters in the System of Spatial Development: Foreign Experience. *Spat. Econ.* 16, 113–135. <https://doi.org/10.14530/se.2020.4.113-135>
- Ketels, C.H.M., 2017. *Cluster Mapping as a Tool for Development*. Institute for Strategy and Competitiveness-Harvard Business School: Boston, MA, USA, 52.

- Koterov, A., Ushenkova, L., Zubenkova, E., Kalinina, M., Biryukov, A., Lastochkina, E., Molodtsova, D., Vaynson, A., 2019. Strength of Association. Report 2. Graduations of Correlation Size. *Med. Radiol. Radiat. Saf.* 64, 12–24. <https://doi.org/10.12737/1024-6177-2019-64-6-12-24>
- Kozhevnikov, S.A., 2019. Spatial and Territorial Development of the European North: Trends and Priorities of Transformation. *Econ. Soc. Chang. Facts, Trends, Forecast* 6, 91–109. <https://doi.org/10.15838/esc.2019.6.66.5>
- Kozyrev, A.N., 2018. Quantum Economics and Quantum Computing in Economics. *Digital Economy* 3, 5–12.
- Kudryavtseva, T.Y., Skhvediani, A.E., 2020a. An econometric analysis of the regional industrial specialization: The Russian manufacturing industry case study. *Econ. Anal. Theory Pract.* 19, 1765–1790. <https://doi.org/10.24891/ea.19.9.1765>
- Kudryavtseva, T.Y., Skhvediani, A.E., 2020b. Studying Regional Clusters with the Use of Data Processing Systems: The Case of the Biopharmaceutical Cluster. *Regionology* 28, 48–79. <https://doi.org/10.15507/2413-1407.110.028.202001.048-079>
- Martínez-Martínez, I., 2014. A connection between quantum decision theory and quantum games: The Hamiltonian of Strategic Interaction. *J. Math. Psychol.* 58, 33–44. <https://doi.org/10.1016/j.jmp.2013.12.004>
- Minakir, P.A., 2017. Theoretical Aspects of the Study of Spatial Economic Systems. *Journal of Economic Theory* 3, 7–10.
- Minsky, M., Papert, S., 1969. *Perceptrons: An Introduction to Computational Geometry*. M.I.T. Press, London.
- Moreira, C., Tiwari, P., Pandey, H.M., Bruza, P., Wichert, A., 2020. Quantum-like influence diagrams for decision-making. *Neural Networks* 132, 190–210. <https://doi.org/10.1016/j.neunet.2020.07.009>
- Pashintseva, N.I., 2018. Methodological Problems of Accounting and Statistics of Research, Development and Technological Works. *Questions of Statistics* 25, 66–72.
- Pozdnyakova, U.A., Popkova, E.G., Kuzlaeva, I.M., Lisova, O.M., Saveleva, N.A., 2017. Strategic Management of Clustering Policy During Provision of Sustainable Development, in: *Integration and Clustering for Sustainable Economic Growth*. Springer, Cham, pp. 413–421. https://doi.org/10.1007/978-3-319-45462-7_40
- Rentkova, K., 2019. The Clusters Phenomenon and Sustainable Regional Development. *IOP Conf. Ser. Mater. Sci. Eng.* 471. <https://doi.org/10.1088/1757-899X/471/10/102039>
- Rodina, V.V., 2014. Comparative Analysis of R&D Financing Mechanisms on the Example of Russia and the United States. *Monitoring Law Enforcement* 4, 65–73.
- Rosenblatt, F., 1962. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books, Washington, D.C.
- Schumpeter, J.A., 1980. *The Theory of Economic Development*. Routledge, London.
- Seber, G.A.F., 1977. *Linear regression analysis*. John Wiley and Sons, New York, N.Y.
- Seidl da Fonseca, R., Pinheiro-Veloso, A., 2018. The Practice and Future of Financing Science, Technology, and Innovation. *Foresight STI Gov.* 12, 6–22. <https://doi.org/10.17323/2500-2597.2018.2.6.22>
- Алферьев, Д.А., 2018. Прогнозирование индикаторов научно-технологического пространства посредством пространственных эконометрических моделей на примере Республики Беларусь. *Белорусская Наука в Условиях Модернизации: Материалы Междунар. Науч.-Практ. Конф., г. Минск, 20-21 Сент. 2018 Г. СтройМедиаПроект, Минск*, –118.
- Егоренко, С.Н., Бондаренко, К.А., Соловьева, С.В., 2018. Инновации: международные сопоставления. Доклад о Человеческом Развитии в Российской Федерации За 2018 Год. Аналитический центр при Правительстве Российской Федерации, Москва, с. 100–123.
- Жуковская, И.Ф., Ивлиева, Н.Н., Труфанова, С.А., 2021. Анализ прямых и косвенных методов стимулирования исследований и разработок в РФ. *Проблемы теории и практики управления* 1, 129–147.
- Козырев, А.Н., 2018. Квантовая экономика и квантовые вычисления в экономике. *Цифровая экономика* 3, 5–12.
- Луговая, Е.С., 2012. Инновации как основа модернизации современного общества. *Вестник Волгоградского университета. Серия 7. Философия. Социология и социальные технологии* 2, 103–108.
- Маслова, Т.С., Лалаева, А.А., 2018. Сравнительный анализ финансирования НИОКР в России и за рубежом. *Бухгалтерский учёт в бюджетных и некоммерческих организациях* 7.
- Минакир, П.А., 2017. Теоретические аспекты исследования пространственных экономических систем. *Журнал экономической теории* 3, 7–10.
- Пашинцева, Н.И., 2018. Методологические проблемы учета и статистики научно-исследовательских, опытно-конструкторских и технологических работ. *Вопросы статистики* 25, 66–72.
- Родина, В.В., 2014. Сравнительный анализ механизмов финансирования НИОКР на примере России и США. *Мониторинг правоприменения* 4, 65–73.
- Степанова, Ю.Н., Лесникова, М.С., 2017. Роль инноваций в современном развитии российского общества. *Международный студенческий научный вестник* 6.
- Ускова, Т.В., 2009. *Управление устойчивым развитием региона: монография*. ИСЭРТ РАН, Вологда.

- Шамин, Р.В., 2019. Машинное обучение в задачах экономики. Грин Принт, Москва.
- Шамин, Р.В., и др. Лекция 5. Обучение перцептрона в задачах распознавания движения.
- Юрченко, Н.Ю., 2013. Финансирование научно-исследовательских и опытно-конструкторских работ в России и за рубежом. Гуманитарный вестник МГТУ им. Н.Э. Баумана 1. <https://doi.org/10.18698/2306-8477-2013-1-33>

The article was submitted 11.04.2021, approved after reviewing 9.06.2021, accepted for publication 17.06.2021.

Статья поступила в редакцию 11.04.2021, одобрена после рецензирования 9.06.2021, принята к публикации 17.06.2021.

About the authors:

1. Rodionov Dmitrii Grigorievich, Doctor of Economics, professor, Head of Higher School of Engineering and Economics, Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia
<https://orcid.org/0000-0002-1254-0464>, drodionov@spbstu.ru
2. Alferyev Dmitrii Alexandrovich, PhD, senior lecturer, Vologda Research Center of the Russian Academy of Sciences (VolRC RAS), Vologda, Russia,
<https://orcid.org/0000-0003-3511-7228>, alferev_1991@mail.ru
3. Klimova Yuliya Olegovna, associate scientist, Vologda Research Center of the Russian Academy of Sciences (VolRC RAS), Vologda, Russia.
<https://orcid.org/0000-0002-3295-9510>, j.uschakowa2017@yandex.ru
4. Alpysbayev Kaisar Serikuly, Candidate of Economic Sciences, Senior Lecturer, Department of Economics and Business, "Kainar" Academy, Republic of Kazakhstan,
<http://orcid.org/0000-0003-3349-701X>, kaisaralp@gmail.com

Информация об авторах:

1. Дмитрий Григорьевич Родионов, д.э.н., профессор, директор Высшей инженерно-экономической школы, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия, <https://orcid.org/0000-0002-1254-0464>, drodionov@spbstu.ru
2. Дмитрий Александрович Алферьев, к.э.н., ассистент, Высшая инженерно-экономическая школа, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия; Научный сотрудник, Вологодский научный центр РАН, Вологда, Россия,
<https://orcid.org/0000-0003-3511-7228>, alferev_1991@mail.ru
3. Юлия Олеговна Климова, младший научный сотрудник, Вологодский научный центр РАН, Вологда,
<https://orcid.org/0000-0002-3295-9510>, j.uschakowa2017@yandex.ru
4. Кайсар Алпысбаев, к.э.н., старший преподаватель, Международный Университет Информационных Технологий, Алматы, Казахстан,
<http://orcid.org/0000-0003-3349-701X>, kaisaralp@gmail.com