The Digital Diversity in Russian Regional Dynamics: Analysis by Machine Learning Methods

Lyudmila Gadasina, Sergey Voitenko and Pasi Luukka

Abstract
The paper focuses on the digital economy of the Russian Federation by analyzing the level of digitization of its regions based on official statistical data from open sources using machine learning methods with verification of the most strongly influencing factors. Hierarchical clustering is applied to determine different groups of regions. Random Forest classification algorithm enabled us to explain the peculiar properties of the different regional groups.

Keywords:
Russian regions, Federal State Statistics, cluster analysis, Random Forest algorithm, digitalization

Lyudmila Gadasina is an Associate Professor of Information Systems at Saint Petersburg State University, Russia.
Sergey Voitenko is an Associate Professor of Information Systems at Saint Petersburg State University, Russia.
Pasi Luukka is a Professor at LUT University, School of Business and Management, Finland.

1 Introduction

The global digitalization of the world creates entirely new problems and challenges: global information interdependence and a sharp increase in the speed of technological, informational, and political-social changes that are defining the economy. In order to accept these challenges and problems, it is necessary to develop modern digitalization tools both at the global and at the country level within its regions. This task is most difficult in large multinational countries, such as Russia, where regions differ significantly in the number of digital devices per capita, the development of information and communication infrastructure, services of e-government and e-commerce.

To progress the digitalization several have developed programs in the digital economy. The program "Digital Economy of the Russian Federation", approved in July 2017, provides for the economic development in five areas: "Information infrastructure", "Information security", "Regulatory regulation", "Formation of research competencies and technological reserves" and "Personnel and education" (Programma 2017). In the same year, the Federal state statistics service issued an order (Prikaz 2018) to support this program by providing information gathering and systematization from various Russian organizations. It includes the preparation of annual reports containing "information on the development and (or) use of advanced production technologies"; "information on the use of information and communication technologies and the production of computer equipment, software, and services in these areas"; "information on the implementation of research and development" and "information on innovative activities of the organization" (Prikaz 2018). Many countries collect and publish similar information for their economy (see e.g., Digital Economy Monitoring 2018).

This paper is devoted to the analysis of the development of Russian Federation regions in the digital economy in the direction of "Information infrastructure" analysis which is based on open data collected on the Federal state statistics service portal in the section "Monitoring the development of information society in the Russian Federation" (Monitoring 2018). Regional development of regions is an important component of the whole country progress. Through the prism of levels of regional development, it is possible to analyze the possibilities and problems of the country as a whole and also to predict the direction of its development.

The Russian state collects a large number of indicators of digital development, but they only show the general situation and do not allow revealing the features of the development of individual regions of the country. In this paper, we decided to apply a combination of several machine learning methods to all collected indicators as raw data. This approach made it possible not only to more accurately assess the development features of each region but also, as a result of tuning the collected indicators themselves, to obtain a more accurate interpretation of the results.

In a number of works, e.g. in Abdrakhmanova et al. (2018) for Russia and Ragnedda and Muschert (2013) for Russia and other countries, a statistical exploratory analysis of the development of digitalization was carried out on the basis of only the aggregated basic indicators of the country, without affecting the regional levels.

Quite a lot of works are devoted to the development of the regions (Kuznetsov & Perova 2014; Kuznetsov et al. 2017; Perova & Neznakomtseva 2016). Kuznetsov and Perova (2014) investigated the factors of development of the digital economy of Russia and neural network models of the dynamics of regional development. Eliseeva et al. (2016) used cluster analysis of innovative activity of donor regions of modern Russia for different periods. Cluster migration was traced and the most significant factors that influenced cluster analysis were identified. Eliseeva et al. (2016) assessed the innovation potential and the main directions of development in the northern regions of Russia by using cluster analysis together with principal component analysis (PCA). In their work, PCA was used to reduce the number of indicators used in cluster analysis. Verzilin et al. (2017) explored the socio-economic activity of Russian Federation regions using open-source data collected from federal and regional authorities, social media and social networks.

Besides the above, a number of studies exist where digitalization is examined and compared on different levels and different entities (unions, countries, regions, etc.). Indicators are usually aggregated to produce meaningful information (Cámara & Tuesta 2017; "Digital Russia" Index 2018; The ICT Development Index 2017). Typically, the aim of these studies is to produce an index that can be used as a one-dimensional measure for digitalization. In the EU, the Digital Economy and Society Index (DESI) is commonly used to compare countries and regions. DESI is a composite index that summarizes relevant indicators on Europe's digital performance across five main dimensions: Connectivity, Human Capital, Use of Internet, Integration of Digital Technology and Digital Public Services[1].

Some papers have used linear regression models to explain determinants of ICT adoption in different groups of countries (see e.g., Billon et al., 2010). Billon et al. (2010) confirmed that there is a positive and significant relationship between GDP and the digitalization index and showed the existence of a positive and significant relationship between the digitalization index and telephone mainlines.

In this paper, we compare the regions of Russia by analyzing the full group of official indicators by using machine learning methods. This differs from earlier research also from the fact that we are not using any aggregated indicators.

2 Methodology and Data Sources

2.1 A methodology of investigation

We apply machine learning techniques to analyze the collected data. At the first stage, several region clusters of the Russian Federation with similar values of a number of indicators for 2014 and 2017 were identified by using hierarchical cluster analysis (Rokach & Maimon 2005). The quality of clusters was checked using a dendrogram analysis. In dendogram analysis, a tree is constructed using a matrix of proximity measures, where each branch of a tree corresponds to indicator.

In the hierarchical clustering, Ward's method (also known as Ward's minimum variance method) was used to merge clusters. For the actual distance in the Ward's method, we selected Euclidean distance. Applying the complete-linkage method gives very close results. Applying single-linkage, unweighted average linkage and weighted average linkage clustering does not lead to forming clusters.

Cluster numbers from hierarchical clustering results in the regional clusters used for the classification task as variables to predict. For the classification method, we used Random Forest (Breiman 2001) classifier, which is an ensemble classifier and can also determine the most influential variables for the prediction task. At the next stage, the selected groups of regions were compared with the rating of their socio-economic development over the same period and for the previous ones. At the last stage, the dynamics of changes in indicators for the period under review were analyzed.

[1] https://ec.europa.eu/digital-single-market/en/desi

2.2 Data Sources

For this study, we used open-access data from the Federal State Statistics Service (Monitoring 2018) related to the digitization of the regions of the Russian Federation in the following areas: the level of digitalization of companies, the use of the Internet by the population, including the use of public services and education, etc. Since for many of the indicators of interest to our study, data has been published since 2014, and data for 2018 were not fully collected at the time of the study, it was decided to limit the period to 2014-2017. The research also used the Rating of socio-economic status of the Russian Federation regions, compiled annually by the Rating Agency "RIA Rating" (Rating 2017).

2.3 Methodology of data collecting

Thirty-three of the following indicators were selected for the study. Table 1 shows the indicator groups and their ranges. The full list of indicators is in Appendix A.

Table 1. Data descriptions

GROUPS OF THE INDICATORS	INDICATORS RANGE IN THE GROUP
The use of computers, computer nets and the Internet in companies.	I1 – I14
Use of special software and information systems in companies.	I15 – I23
Education: technology support.	I24 – I27
Internet use by the public.	I28 – I33

In order to eliminate strongly dependent variables, an analysis of their interrelations was carried out in order to select statistically independent indicators. For a pairwise comparison of all 33 indicators, correlation coefficients were calculated and scattering clouds were constructed. Table 2 shows the highest correlation coefficients for indicators.

Table 2. Correlation coefficients for indicators.

INDICATOR NUMBER	INDICATOR NUMBER	CORRELATION
I1	I7	0.891
I1	I11	0.873
I1	I12	0.845
I2	I22	0.820
I4	I18	0.820
I5	I16	0.837
I6	I8	0.894
I7	I12	0.904
I11	I12	0.895
I15	I22	0.836
I16	I17	0.889
I16	I18	0.891
I16	I23	0.895
I17	I18	0.875
I28	I29	0.823

In order to get rid of highly correlated indicators, we selected only those which had higher variance. In the cases where the variances differed insignificantly, the final set includes indicators with a better interpretation. Thus, indicators: I5, I6, I7, I11, I12, I18, I22, I23 & I28 were excluded from the final set of 24 indicators.

3 Results

3.1 2017 indicators

Clustering analysis using hierarchical clustering with the Ward method for merging clusters was performed for the data using the selected 24 indicators. For the number of clusters, we found five to be the most suitable for grouping different regions. Figure 1 shows dendrogram with marked 5 clusters.

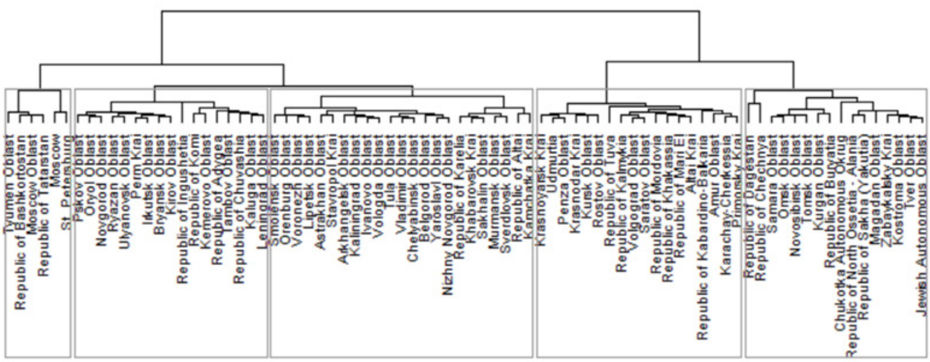

Fig. 1. Dendrogram with 5 clusters shown

The next step was to apply Random Forest to classify our data set. Random Forest can also identify the most significant indicators from the classification point of view by using mean accuracy decrease. In this study, we found that most significant indicators were: I9, I8, I2, I20, I10, and I13 (in descending order of importance).

The Random Forest algorithm parameters were selected and a series of classification experiments were carried out with the initial set of parameters, 24 selected parameters, and 6 most significant ones. Table 3 shows the out-of-bag (OOB) error for the algorithm, which is the average error by algorithm applying to the sample do not contain in the Forest respective bootstrap sample.

Table 3. Random Forest machine learning algorithm

NUMBER OF USING FEATURES	MEAN ERROR	STANDARD DEVIATION FOR ERRORS
Full dataset of 33 features	16,4%	1,48
24 features	17,7%	1,19
6 most important features	31,5%	1,2

Table 3 shows than using only the most important indicators is not enough for clusters predicting. Other indicators can give some special characteristics for clusters.

List of identified features shows that the companies in different regions most significantly differ in the following indicators:

- 'use of Internet access at a speed of at least 2 Mbit';
- 'percentage of organizations using broadband Internet access in the total number of organizations';
- 'using local area networks by organizations';
- 'using electronic data interchange between organizations own and external information systems';
- 'availability of a website';
- 'provision of technical means for mobile Internet access to its employees'.

Figure 2 allows comparing indicator I9 (Percentage of organizations using Internet access at a speed of at least 2 Mbit/s in the total number of organizations) with the help of boxplots for five clusters. It clearly shows the dissimilarities of this indicator value probability distributions for different clusters.

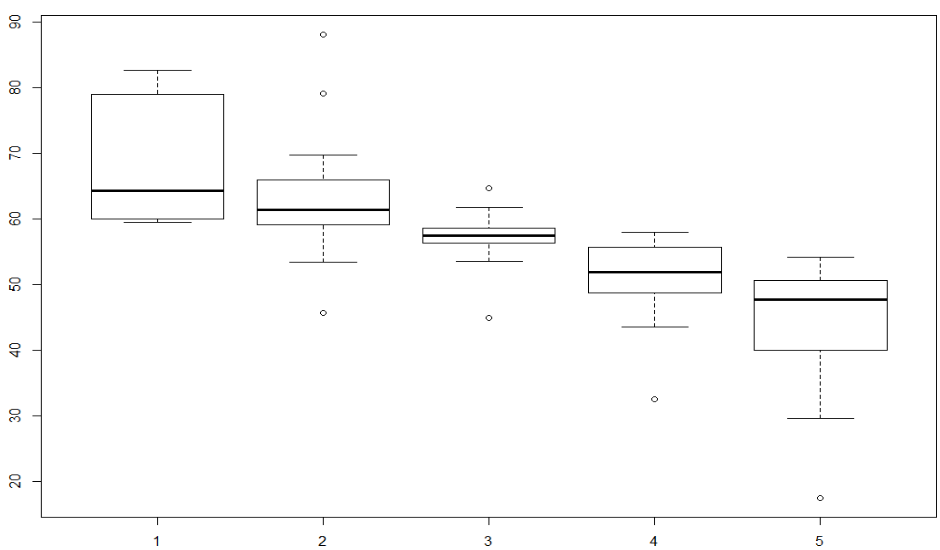

Fig. 2. Compare Indicator I9 boxplots for five clusters.

The least significant indicators were: I24, I27, I25 (in descending order of importance). This means that indicators such as 'number of personal computers used for educational purposes in educational institutions', 'the proportion of students enrolled in training programs for skilled workers', 'employees and percentage of educational institutions of higher vocal education connected to the Internet in the total number of institutions of higher vocal education' – varies slightly from region to region. All these indicators belong to the field of education, and their average values for all regions are 12.88, 0.42 and 97.25, respectively.

Table 4 provides estimates of indicators mean values for each cluster.

Table 4. 2017 indicators mean values for five clusters.

INDICATOR	CLUSTER 1	CLUSTER2	CLUSTER 3	CLUSTER 4	CLUSTER 5
I1	53.17	47.22	45.5	47.24	39
I2	68.32	64.95	59.99	49.11	36.5
I3	34.97	26.89	22.96	17.68	7.4
I4	22.32	16.65	13.84	12.81	6.2
I8	89.18	86.43	83.25	73.52	74.5
I9	68.28	59.95	54.39	47.22	40.3
I10	56.28	50.72	44.32	38.84	41.9
I13	46.4	35.65	29.12	27.63	19.85
I14	5.12	2.74	2.41	2.14	1.05
I15	40.73	37.59	34.67	30.72	15.55
I16	29.32	21.49	18.62	16.03	5.35
I17	18.6	11.21	9.48	7.82	2.2
I19	70	69.41	69.13	61.29	41.6
I20	68.93	64.79	65.34	56.59	40.9
I21	6.52	4.52	3.91	3.94	1.45
I24	17.17	13.92	11.7	11.18	7
I25	94.2	96.54	99.33	97.64	95.45
I26	3.485	2.35	2.17	2.91	2.11
I27	0.33	0.42	0.44	0.42	0.35
I29	79.22	72.45	70.29	73.71	61.8
I30	61.75	39.33	23.65	35.49	28.8
I31	74.67	62.62	47.41	58.76	46.15
I32	80.13	71.31	64.99	73.36	50.9
I33	37.67	28.67	22.49	24.99	10.3

Clustering analysis results show that regions can be ranked across the entire spectrum of selected indicators, with the exception of indicators I25 and I27. The ranking order is determined by the number of the cluster, which contains the region in terms of digitizing the selected indicators.

The *first cluster* consists of the following regions: Moscow, Moscow obl., St. Petersburg, Tatarstan, Tyumen obl., Bashkortostan.

The *second cluster* consists of 37 regions, which are the following:

Regions (oblasts): Arkhangelsk, Astrakhan, Belgorod, Chelyabinsk, Ivanovo, Kaliningrad, Kaluga, Kamchatka, Kemerovo, Khabarovsk, Kursk, Leningrad, Lipetsk, Murmansk, Nizhny Novgorod, Orenburg, Penza, Rostov, Sakhalin, Smolensk, Sverdlovsk, Tambov, Tula, Vladimir, Vologda, Voronezh, Yaroslavl; provinces: Krasnodar, Krasnoyarsk, Primorsk, Stavropol; republics: Adygea, Altai, Chuvash, Karelia, Khakassia, Udmurtia.

The *third cluster* consists of 20 regions, which are the following:

Regions (oblasts): Amur, Bryansk, Irkutsk, Kirov, Kostroma, Magadan, Novgorod, Orlovsky, Pskovskaya, Ryazan, Tver, Ulyanovsk; provinces: Perm, Zabaykalsky, Chukotka Autonomous

Okrug, Jewish Autonomous; republics: Ingushetia, Komi, Mari El.

The *fourth cluster* consists of 17 regions, which are the following:

Regions (oblasts): Kurgan, Novosibirsk, Omsk, Samara, Saratov, Tomsk, Volgograd; provinces: Altai; republics: Buryatia, Mordovia, Kabarda-but-Balkar, Kalmykia, Karachay-Cherkessia, North Ossetia-Alania, Sakha (Yakutia), Tuva.

The *fifth cluster* consists of 2 regions: the Republic of Dagestan, the Republic of Chechnya.

Comparing with the Rating of the socio-economic situation of the subjects of the Russian Federation in 2017 (this rating is formed annually by "Rating Agency" RIA Rating ", (Rating 2017)), we can conclude that the first cluster of leaders included regions with a rating of not less than 11 (also on the ratings since 2014). This means that in regions with high socio-economic indicators, companies have more opportunities for digitalization. For other clusters, there is no clear link between the level of socio-economic development and the level of digital development.

3.2 2014 indicators

Cluster analysis of data for 2014 also allowed forming five well-distinguishable groups of regions. Similarly, in 2017, a group of leaders stood out, as well as two groups of leaders (first and second place) and the group with the lowest indicators. But clusters number 3 and 4 were formed according to the following principle:

- in one cluster, regions with the lowest developed indicators of the category *Internet use by the public* with average indicators of other spheres;
- in other regions with the lowest indicators of the category *Use of special software and information systems in companies* and several indicators of the *Use of computers, computer nets and the Internet in companies* category.

The composition of the clusters themselves also differs from 2017 year clusters. For example, the cluster of leaders consisted only of Moscow, St. Petersburg. Cluster number 5 contains 17 regions.

3.2 Changes in the indicators over time

The next stage of the study is devoted to the changes in the values of indicators, which are calculated as a relative increment of indicators in 2017 relative to the values of 2014 ($k = 1,2,3,\dots 33$):

$$\Delta Ik = \frac{\left(Ik(2017) - Ik(2014)\right)}{Ik(2014)} \quad \Delta Ik = \frac{\left(Ik(2017) - Ik(2014)\right)}{Ik(2014)} \quad (1)$$

After analyzing the changes, we identified three clusters and three separate regions: the Republic of Dagestan, the Republic of Ingushetia, the Republic of Chechnya. Figure 3 shows dendrogram with 3 clusters, three separate regions are missed. Figure 4 shows the clustering heights – the value of the criterion associated with the Ward's method of clustering. These two figures give us the possibility to choose 3 or 7 clusters, but 7 clusters are not clear for its characteristics interpretation.

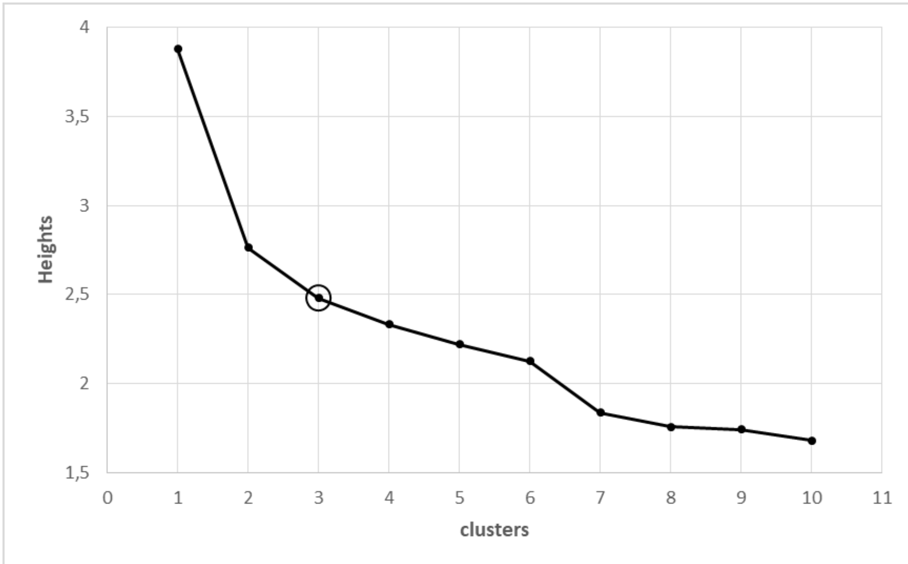

Fig. 4. Elbow diagram

The characteristics of the average values of the indicators included in each cluster are presented in table 5.

Table 5. Indicators mean values for four clusters.

INDICATOR	CLUSTER 1	CLUSTER2	CLUSTER 3
ΔI1	0.07	0.09	0.06
ΔI2	-0.14	-0.01	-0.05
ΔI3	0.52	0.68	0.75
ΔI4	0.16	0.25	0.33
ΔI8	-0.02	0.09	0.05
ΔI9	0.07	0.21	0.26
ΔI10	0.09	0.31	0.24
ΔI13	0.26	0.73	0.36
ΔI14	0.85	1.28	1.31
ΔI15	-0.07	0.16	0.06
ΔI16	0.02	0.19	0.19
ΔI17	0.16	0.08	0.38
ΔI19	0.07	0.27	0.15
ΔI20	0.12	0.44	0.21
ΔI21	0.11	0.09	0.39
ΔI24	0.03	0.16	0.05
ΔI25	0.02	0.05	0.04
ΔI26	-0.19	-0.21	-0.27
ΔI27	-0.18	-0.13	-0.30
ΔI29	0.12	0.26	0.16
ΔI30	3.79	10.66	5.30
ΔI31	0.90	2.84	1.37
ΔI32	0.10	0.30	0.23
ΔI33	0.67	1.57	0.86

In general, the greatest changes occurred on indicators:

- Percentage of workers using mobile Internet access provided by an organization at least once a week in the total number of payroll organizations;
- Proportion of the population that used the Internet to obtain state and municipal services in the total population surveyed,
- The share of the population that used the Internet to obtain state and municipal services in the total population that received state and municipal services.

These changes in the last two parameters are associated with the introduction of the Program E-government in Russia.

Also, the greatest change occurred in the “backward” regions according to the studied set of parameters. In particular, in regions with high subsidies and investments, such as the Republic of Dagestan, the Republic of Ingushetia, the Republic of Chechnya.

The smallest changes occurred in the following regions: Moscow, St. Petersburg, Arkhangelsk oblast, Chelyabinsk oblast, Irkutsk oblast, Irkutsk oblast, Kemerovo oblast, Krasnoyarsk oblast, Novosibirsk oblast, Omsk oblast, Perm oblast, Samara oblast, Saratov oblast, Sverdlovsk oblast, Tomsk oblast, Tyumen oblast, Volgograd oblast, Republic of Adygea, Republic of Buryatia, Republic of Karelia, Republic of Khakassia, Republic of Komi, Republic of Mari El, Republic of Tatarstan, Udmurt Republic, Primorsky Krai, Khabarovsk Krai.

4 Conclusion

This paper presents a comparative analysis of the digital development of regions based on indicators from open sources. The results allow us to conclude that the digital economy of Russia can be characterized as “digital inequality” due to disproportion in the IT development of regions. Lead regions (included in cluster number 1) can be compared in their level of digitalization with world leaders, while other regions are more comparable with the states of the Active Followers and Lagging Followers categories. The digital inequality generated by the existing economic and social gaps between the capital and the other regions contains in itself the possibilities to overcome it due to the rapid and relatively inexpensive scaling inherent in digital solutions and services.

It should be noted that the so-called “lagging” regions showed a significant improvement in year 2017 compared to year 2014, which is largely due to state support. Because the national program “Digital Economy in Russian Federation” was approved in 2017, it makes sense to conduct a detailed monitoring of the development of the regions and changes in their position for the targeted direction of the state’s attention in these regions, which will contribute to a more uniform development of the country in the digital economy.

References

Abdrakhmanova, G., Vishnevskiy, K., Volkova, G., Gokhberg, L. et al. (2018). *Digital Economy Indicators in the Russian Federation: 2018: Data Book*. National Research University Higher School of Economics. Moscow: HSE. (in Russian)

Billon, M., Lera-Lopez, F. & Marco, R. (2010). Differences in digitalization levels: a multivariate analysis studying the global digital divide. *Review of World Economics* 146:1, 39-73.

Breiman, L. (2001). Random forests. *Machine Learning* 45:1, 5-32.

Cámara, N., Tuesta, D. (2017). *DiGiX: The Digitization Index*. BBVA Research.

Digital Economy Monitoring Report 2018. Federal Ministry for Economic Affairs and Energy.

“Digital Russia” Index. Reflection of Russian regions digitalization process through the prism of open sources. Know-how methodology that takes into consideration Russian specifics and best practices. (2018). Moscow School of Management SKOLKOVO.

Eliseeva, I.I., Trot, K.S. & Mukhametzyanova, F.G. (2016). Cluster analysis of modern Russian donorregion. *Herald of the Belgorod University of Cooperation, Economics and Law* 2(58), 117-125 (in Russian).

Gadzhiev, Yu.A., Styrov, M.M., Kolechkov, D.V. & Shlyakhtina, N.V. (2016). Analysis of innovation potential of Northern Russian regions. *Economic and Social Changes: Facts, Trends, Forecast* 6, 236-254.

Kuznetsov, Yu.A., Perova, V.I. (2014). Neuronetwork modeling of dynamics of innovative development of the regions of the Russian Federation. *Regional Economics: Theory and Practice* 4:331, 18-28 (in Russian).

Kuznetsov, Yu.A., Perova, V.I., Semikov, D.S. (2017). Information and communication technologies as a factor in the development of digital economy in the Russian Federation. *Vestnik of Lobachevsky University of Nizhni Novgorod* 4:48, 38-47 (in Russian).

Monitoring the development of the information society in the Russian Federation (on 03.10.2018). Digital economy of the Russian Federation.

Perova, V.I., Neznakomtseva, O.Yu. (2016). Research into the dynamics of social and economic development of the Russian Federation regions. *Vestnik of Lobachevsky University of Nizhni Novgorod* 4:44, 44-51 (in Russian).

Prikaz Rosstata ot 30.08.2017 № 563 (red. ot 28.03.2018) "Ob utverzhdenii statisticheskogo instrumentariya dlya organizacii federal'nogo statisticheskogo nablyudeniya za deyatel'nost'yu v sfere obrazovaniya, nauki, innovacij i informacionnyh tekhnologij"

Programma "Cifrovaya ehkonomika Rossijskoj Federacii" [Programme "Digital Economy in Russian Federation"]. Pravitel'stvo Rossii.

Ragnedda, M., Muschert, G. W. (ed.) (2013). *The Digital Divide: The Internet and Social Inequality in International Perspective*. Routledge.

Rating of socio-economic situation of the Russian Federation. The end of 2017. The rating Agency "RIA Rating".

Rokach, L., Maimon, O. (2005). *Clustering Methods. Data Mining and Knowledge Discovery Handbook*. Springer.

The ICT Development Index (IDI): conceptual framework and methodology.

Verzilin, D., Maximova, T., & Sokolova, I. (2017). Online socioeconomic activity in Russia: patterns of dynamics and regional diversity. *International Conference on Digital Transformation and Global Society* pp. 55-69. Springer.

Appendix A.

The list of indicators.

The use of computers, computer nets and the Internet in companies.

I1 – The number of personal computers per 100 employees of organizations.

I2 – Percentage of organizations using local area networks in the total number of surveyed organizations.

I3 – Percentage of organizations using the Intranet in the total number of surveyed organizations.

I4 – Percentage of organizations using Extranet in the total number of surveyed organizations.

I5 – Percentage of organizations using open-source operating systems provided by third parties (for example, Linux) in the total number of organizations surveyed.

I6 – Percentage of organizations using the Internet in the total number of surveyed organizations.

I7 – Number of personal computers with access to the Internet per 100 employees of organizations.

I8 – Percentage of organizations using broadband Internet access in the total number of organizations.

I9 – Percentage of organizations using Internet access at a speed of at least 2 Mbit / s in the total number of organizations.

I10 – Percentage of organizations that had a website in the total number of surveyed organizations.

I11 – Percentage of employees of organizations that used personal computers at least 1 time per week, in the total number of payroll organizations.

I12 – Percentage of employees of organizations that used the Internet at least 1 time per week, in the total number of payroll organizations.

I13 – Percentage of organizations that provided technical means for mobile Internet access to their employees, in the total number of surveyed organizations.

I14 – Percentage of workers using mobile Internet access provided by an organization at least once a week in the total number of payroll organizations.

Use of special software and information systems in companies.

I15 – Percentage of organizations that had special software for managing the procurement of goods (works, services) in the total number of organizations surveyed.

I16 – Percentage of organizations that had special software for managing the sales of goods (works, services) in the total number of organizations surveyed.

I17 – Percentage of organizations using ERP systems in the total number of surveyed organizations.

I18 – Percentage of organizations using CRM systems in the total number of surveyed organizations.

I19 – Percentage of organizations using electronic document management systems in the total number of surveyed organizations.

I20 – Percentage of organizations that used electronic data interchange between their own and external information systems on exchange formats (EDIFACT, EANCOM, ANSI X12; XML-based standards, such as ebXML, RosettaNet, UBL, papiNET; proprietary standards agreed between organizations) in the total number of surveyed of organizations.

I21 – Percentage of organizations that used SCM-systems in the total number of surveyed organizations.

I22 – Percentage of organizations that placed orders for goods (works, services) on the Internet, in the total number of surveyed organizations.

I23 – Percentage of organizations that received orders for manufactured goods (works, services) over the Internet, in the total number of surveyed organizations.

Education: technology support.

I24 – Number of personal computers used for educational purposes per 100 students of state and municipal educational institutions.

I25 – Percentage of educational institutions of higher vocational education connected to the Internet in the total number of institutions of higher vocational education surveyed.

I26 – The number of students enrolled in educational programs of higher education - undergraduate programs, specialties, graduate programs, per 100 population.

I27 – The proportion of students enrolled in training programs for skilled workers, employees, in the total population.

Internet use by the public.

I28 – Number of Internet users per 100 population.

I29 – The proportion of the population that is active users of the Internet in the total population.

I30 – Proportion of the population that used the Internet to obtain state and municipal services in the total population surveyed.

I31 – The share of the population that used the Internet to obtain state and municipal services in the total population that received state and municipal services.

I32 – Proportion of households with broadband Internet access in the total number of households.

I33 – The share of the population that used the Internet to order goods and (or) services in the total population.