http://www.moderntechno.de/index.php/meit/article/view/meit41-01-002
DOI: 10.30890/2567-5273.2025-41-01-002

UDC 004.42:004.855.5:519.237.8

SUPPORTING DECISION-MAKING IN THE SEGMENTATION OF TELECOMMUNICATIONS COMPANY CUSTOMERS USING SPECIALIZED SOFTWARE

Oleksandr Ivashchenko
PhD student
ORCID: 0009-0007-5470-9137
Serhii Fedin
d.t.s., prof.
ORCID: 0000-0001-9732-632X
National Transport University, Ukraine
01010, Kyiv, M. Omelianovych-Pavlenko str. 1

Abstract. *The article presents the results of segmentation of the customer database of a telecommunications company, performed on the basis of cluster analysis of business data using the Kohonen self-organizing maps method. To increase the efficiency of the analysis of customer behavioral characteristics, a specialized software application has been developed that allows automatic clustering, visualization of a map of customer segments, and calculation of statistical characteristics of selected customer groups. A comparative analysis was conducted between the developed application and the Kohonen maps module used in the STATISTICA statistical analysis system. The developed software application, unlike the Cluster analysis module of the STATISTICA system, allows increasing the efficiency of the analytical calculation process by the proposed means of visualization and processing of the obtained results. The results demonstrate that the application can be employed to facilitate decision-making in customer base segmentation and the creation of targeted marketing initiatives in the telecommunications field.*

Key words: *segmentation, clustering, Kohonen self-organizing map, software application, STATISTICA.*

Introduction

In the current conditions of the telecommunications market development, one of the key tasks for operator companies is effective management of interaction with customers. A large number of subscribers, high competition, and constantly growing consumer expectations require the implementation of analytical tools that allow not only to analyze existing data but also to form personalized offers based on it. In this context, customer base segmentation becomes a strategic tool for decision support (DS) in assessing marketing effectiveness, increasing loyalty, and optimizing company resources [1].

Since the typical customer base of a telecommunications company is large in volume and multidimensional in structure, it is advisable to use data mining methods to identify new, non-trivial, practically useful, and interpretable knowledge in large data sets [2]. One effective means of data mining is clustering algorithms.

Among the software solutions that implement such algorithms, various statistical analysis systems have become widely used, in particular STATISTICA, which provides tools for building cluster models, including the use of Kohonen's Self-Organizing Maps (SOM). However, the Cluster analysis module of the STATISTICA system has certain limitations. This particularly concerns the need for additional processing of clustering results, the complexity of obtaining statistical characteristics of clusters, and limited visual analysis capabilities.

At the same time, there is a need for more flexible and specialized solutions that can take into account the specifics of the subject area, simplify the process of processing results, and provide better visualization. In this regard, it is important to create a specialized software tool capable of implementing the SOM algorithm to ensure convenient analysis of customer data in the telecommunications sector.

Literature review

Customer base segmentation is one of the key tools of modern marketing analysis and CRM (Customer Relationship Management) systems, enabling more effective customer relationship management, product adaptation based on accumulated information, and increased customer satisfaction [3]. Data on customer behavior, preferences, and interaction history from CRM systems and internal company information sources form the basis for further classification of consumers according to similar characteristics. According to the classic approaches to marketing by Kotler and Keller, segmentation involves dividing customers according to certain characteristics in order to develop targeted communication and service strategies [4].

At present, where customer base segmentation is becoming the basis for developing personalized strategies, the use of decision support systems (DSS) is relevant. According to the definition proposed by Daniel J. Power [5], DSS are interactive computer systems that help users use computer communications, data, documents, knowledge, and models to solve problems and make decisions. These systems combine analytical models, data visualization, and user interfaces to make informed management decisions. In the telecommunications industry, the use of DSS allows not only to segment the customer base, but also to quickly make decisions on optimizing tariff plans, marketing strategies, or managing customer attrition, which is especially important in the context of highly dynamic consumer behavior.

The scientific literature describes a number of types of segmentation: geographic, demographic, psychographic, behavioral, and combined. Behavioral segmentation is particularly relevant in the telecommunications industry, where there are dynamic changes in service consumption, a wide variety of communication channels, and the growing role of digital services. As noted by M. Wedel and W. A. Kamakura [6], the behavioral approach makes it possible to take into account the actual actions of the customer – call frequency, traffic volume, intensity of service use – and, accordingly, to form flexible personalized offers.

Segmentation of the customer base creates the basis for the implementation of targeted marketing strategies, understood as targeting product and communication offers to specific groups of consumers, taking into account their behavioral characteristics and needs [7]. As marketing and sales experts point out, focusing on clearly defined segments makes it possible to improve communication effectiveness, optimize advertising campaign costs, and ensure personalized interaction with customers. The use of such a targeted approach is particularly relevant for the telecommunications industry, as the diversity of telecommunications service consumption models requires flexible and carefully designed marketing strategies.

The methodological basis for segmentation is based on the use of statistical data analysis methods. The most common among them are cluster analysis algorithms, in particular, the k-means method, hierarchical clustering, DBSCAN, as well as dimension reduction methods such as PCA and t-SNE [8]. As emphasized by J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson [9], the choice of a specific clustering algorithm depends on the number of variables, the type of scales, and the goals of further interpretation of the results. In a study by S. V. Oliinyk and B. M. Boiko [10], the effectiveness of k-means, hierarchical clustering, DBSCAN, and Affinity Propagation methods for customer segmentation was compared. The best results were shown by the k-means and Affinity Propagation algorithms, the application of which made it possible to identify clearly defined groups based on cost, purchase frequency, and product range.

According to D. Hand, H. Mannila, and P. Smyth [11], data mining is an interdisciplinary field that combines statistics, machine learning, and database management to discover previously unknown useful information.

Modern approaches to segmentation are increasingly based on a combination of feature reduction methods and deep cluster analysis. In particular, the work of F. J. Farahani and S. Tabibian [12] considers a multi-step architecture for segmenting telecommunications company customers, involving preliminary transformation of input data to improve its structure before applying clustering algorithms. This approach creates conditions for the development of personalized interaction strategies and improves the quality of structured behavioral group identification.

An alternative method is proposed in a study by P. Twomey and R. Cadman [13], in which consumer behavior in telecommunications and media markets is analyzed using agent-based modeling. Such models allow for interpersonal influence, social effects, and adaptive behavior to be taken into account, which makes it possible to more accurately model the dynamics of customer decisions and predict reactions to changes in services.

At the same time, modern research is shifting its focus to combining classic cluster analysis methods with powerful tools for representing consumer behavior, in particular the RFM (Recency, Frequency, Monetary) method. For example, the work of M. Sarkar, A. Roy Puja, and F. R. Chowdhury [14] demonstrates the effectiveness of using the k-means algorithm in conjunction with RFM analysis to form accurate customer clusters, which allows for highly effective personalized targeted marketing strategies. In addition, the work of M. A. Gomes and T. A. Meisen [15] summarizes an approach to personalized customer targeting, which involves segmentation as one of the key stages. The authors identify four main phases of this process: information gathering, customer representation building, cluster analysis, and targeted actions. This multi-stage approach provides flexibility in taking into account changes in consumer behavior and increases the accuracy of customer base segmentation.

In this context, the application of SOM deserves special attention. This method, proposed by Teuvo Kohonen [16], is based on the principles of competitive learning and allows reducing the dimension of the feature space while preserving the topological proximity of the data. Kohonen's algorithm is actively used in customer clustering tasks based on service consumption models. In addition, the method demonstrates high efficiency in tasks of visualizing the cluster structure of data. In the work of T. Schreck, J. Bernard, T. von Landesberger, and J. Kohlhammer [17], SOM was used to analyze customer trajectory data, which allowed the identification of recurring behavior patterns. The effectiveness of Kohonen maps in visualizing and analyzing the spatial distribution of customers is confirmed in the work of A. Ultsch and C. Vetter [18], where SOM is considered as an alternative to traditional clustering methods, in particular when processing large amounts of data while preserving the topological structure.

It is also worth noting the work of R. Wehrens and L. Buydens [19], which presents their kohonen module for the R environment, implementing the Kohonen self-organizing map method. The developed toolkit supports the construction of classical SOMs, as well as extended variants for classification, regression, and integration of different types of data. The authors emphasize the possibilities of visualizing results and applying supervised learning (using pre-classified data), which significantly expands the potential for using SOM in applied research.

Among the ready-made tool solutions, the STATISTICA system has become widely used, combining traditional methods of statistical data analysis, decision trees, neural networks, and visualization tools. As stated in the official documentation, STATISTICA allows you to build interpretable customer segmentation models and integrate the results into the company's business processes [20].

Thus, modern research confirms the effectiveness of using Kohonen self-organizing maps for solving customer segmentation problems, especially in the telecommunications sector. At the same time, existing methods demonstrate limitations related to insufficient integration of the process of analysis, visualization, and interpretation of results in a single environment. This determines the relevance of developing specialized software solutions that simplify the segmentation process, improve the convenience of working with results, and adapt analytical tools to the needs of telecommunications companies.

Research objective and tasks

The purpose of the study is to conduct a comparative analysis of the results of cluster analysis obtained using a specialized software application developed taking into account the criteria of cluster interpretability, ease of visualization and processing of results, and the results obtained in the STATISTICA environment using Kohonen self-organizing maps.

To achieve this goal, it is necessary to:

- perform clustering of the customer base in the created software application and in the STATISTICA 12 statistical data analysis system;
- conduct a comparative assessment of the clustering results based on quantitative and qualitative indicators;
- to evaluate the functional features of the Cluster analysis module of the STATISTICA 12 system and the developed specialized software application in their practical application as decision support tools in the telecommunications sector.

Materials and research methods

Substantive formulation of the problem

Considering that commercial statistical data analysis software systems, including STATISTICA, provide the ability to build self-organizing maps, the process of grouping nodes into clusters and calculating average values of characteristics of observations assigned to these nodes in such systems is not automated and requires additional operations. In particular, in the Cluster analysis module of STATISTICA, it is necessary to perform a clustering propagation procedure, during which each example in the input sample is assigned to the nearest node of the map according to the minimum distance criterion. After that, statistical indicators must be calculated separately for each group. This organization complicates result analysis and increases the time spent working with large data sets.

Taking these features into account, the following requirements were set for the creation of a new software application:

- to implement the construction of an interactive Kohonen map with a more detailed graphical representation of the clustering results compared to the standard capabilities of the STATISTICA system;
- provide the ability to interactively select clusters without the need to perform the clustering propagation procedure;
- ensure automatic calculation of the average values of all behavioral characteristics for selected map nodes.

Input data and training parameters for the Kohonen self-organizing map

The input data for building the SOM model is subscriber data based on information collected from the telecommunications company's billing systems and partially supplemented with data from open sources. Since the sample does not contain personal data that would allow the identification of specific individuals, such information is considered anonymous. The sample includes the following parameters: average monthly mobile communication expenses, number of SMS messages, call activity at different times of the day (morning, afternoon, evening, night), call duration, and the proportion of calls to landline numbers. The combination of these characteristics makes it possible to assess a typical model of user behavior.

The Kohonen self-organizing map model was constructed based on a sample of 4493 observations, each representing a separate subscriber with corresponding behavioral characteristics.

To build a two-dimensional map, the dimension of the neuron grid corresponding to the plane of the self-organizing map nodes was determined.

The dimension of the map was calculated using the empirical formula [21]:

$$m = 5 \cdot \sqrt{n} \tag{1}$$

where n is the number of training samples (n=4493) and m is the total number of neurons on the map. Based on the empirical rule (1), which is widely used in SOM construction, the approximate number of neurons m≈335 is determined. To construct the map, a grid size of 18×24 is selected, which gives 432 neurons, i.e., more than the approximately recommended number. Increasing the number of neurons compared to the calculated value allows for a detailed representation of the data structure, improves segment differentiation, and reduces the likelihood of combining heterogeneous objects within a single node, which is important for improving the interpretability of the resulting clusters.

Neural network training is set to 500 cycles (Training cycles), which provides a sufficient number of iterations to adapt the weight coefficients in each of the neurons. In the created software application, the initial learning rate is set at 0.03 with a gradual decrease in this parameter during training. In the STATISTICA 12 statistical data analysis system, the learning rate is set at 0.1 for the initial stage and 0.02 for the final stage. The differences in learning rate settings are due to the specifics of parameter implementation in different software environments, as well as the desire to ensure a balance between learning speed and stability in the formation of the topological structure of the map in each case.

Segmentation of a telecommunications company's customer base

A specialized software application for segmenting the customer base of a telecommunications company using the Kohonen self-organizing map algorithm has been implemented in the C# programming language on the .NET platform.

Figure 1 shows a fragment of code responsible for initializing the Kohonen self-organizing network map. In each node of the map, initial weight vectors are formed by generating random values within the ranges of minimum and maximum values of the training sample features. The implementation is performed without the use of third-party libraries, which ensures the autonomy of the software solution and the possibility of flexible algorithm configuration.

The developed software application can be considered an example of a specialized DSS focused on customer analytics tasks in the telecommunications sector. Such a system allows analysts and marketers to make informed decisions about tariff personalization, communication strategy optimization, and the identification of priority subscriber groups.

```
/// <summary>
/// Trains a self-organized map using the specified training data.
/// </summary>
1 reference | Oleksandr, 240 days ago | 1 author, 1 change
public virtual void Initialize(IList<DataCollection> models, List<string> columns)
{
    Depth = columns.Count;
    DataDictionary maxValues = new DataDictionary();
    foreach (var column in columns)
    {
        maxValues.Add(column, models.Max(x => x.GetDoubleValue(column)));
    }

    DataDictionary minValues = new DataDictionary();
    foreach (var column in columns)
    {
        minValues.Add(column, models.Min(x => x.GetDoubleValue(column)));
    }

    RandomGenerator random = new RandomGenerator(Seed);
    for (int x = 0; x < Width; x++)
    {
        for (int y = 0; y < Height; y++)
        {
            DataCollection weights = new DataCollection();

            for (int i = 0; i < columns.Count; i++)
            {
                double min = minValues.GetDoubleValue(columns[i]);
                double max = maxValues.GetDoubleValue(columns[i]);
                weights.Add(new DataModel
                {
                    Key = columns[i],
                    Value = (random.Next() * (max - min)) + min,
                    MaxCollectionValue = max
                });
            }

            this[x, y] = new Node(x, y, weights);
        }
    }
}
```

Figure 1 - Fragment of the code for initializing the Kohonen self-organizing map

Neural network training in the application begins after setting the map parameters and starting the training process. During the SOM training process, basic service information about the current epoch, training duration, and processing speed is displayed (Figure 2).

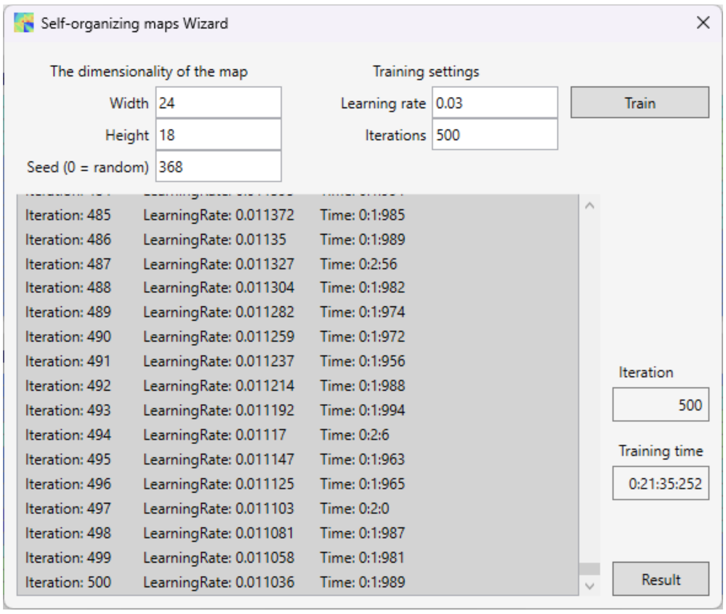


Figure 2 - Setting up and training a Kohonen neural network in the created software application

After training is complete, an interactive window opens with a constructed map, where each node corresponds to a cluster of matched customers. The map is displayed as a two-dimensional grid, on which the density of objects is indicated by a color scale (Figure 3).

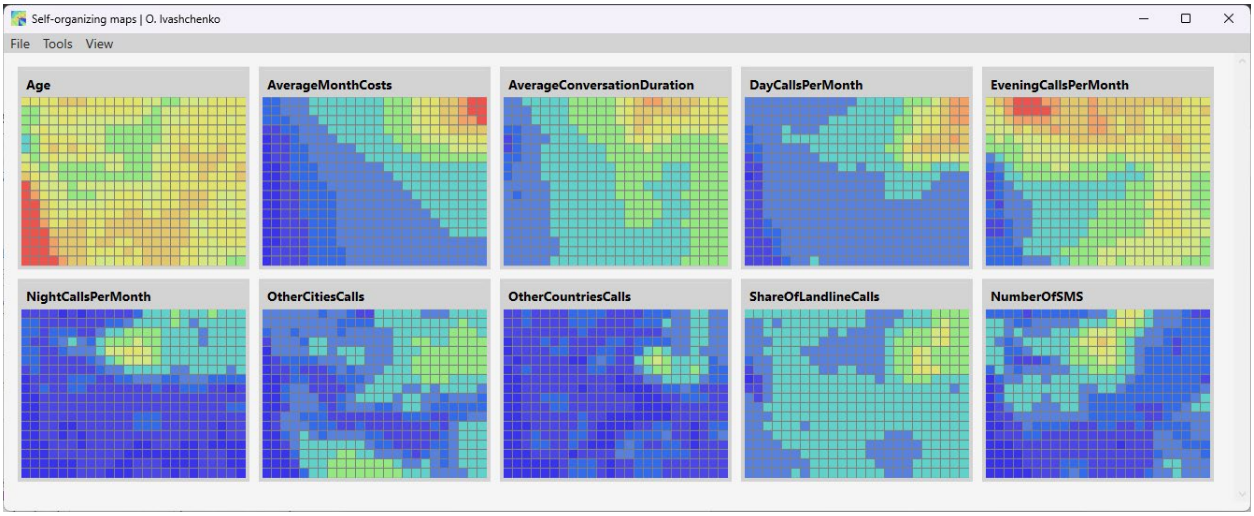


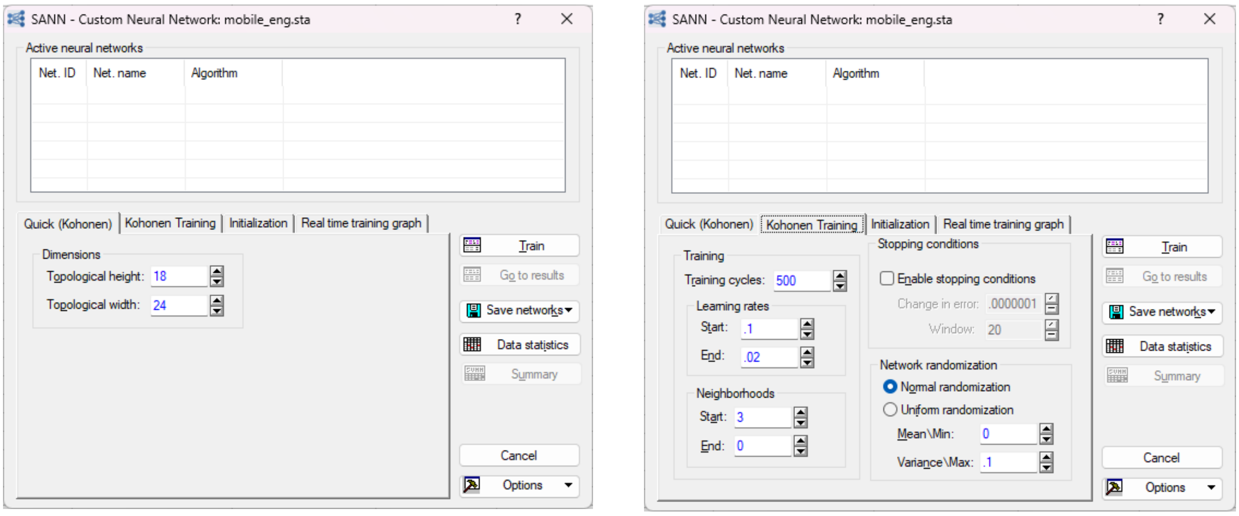
Figure 3 - Visualization of the results of clustering the customer base according to behavioral characteristics in the created software application

The user can select one or more nodes on the map, after which the system automatically calculates the average values of all behavioral characteristics of subscribers belonging to the selected cluster (in particular, expenses, number of calls at different times of the day, SMS activity, etc.) and displays them in a form convenient for analysis. The visual structure of the map makes it easy to identify local clusters of subscribers with similar behavior, which creates convenient conditions for further interpretation of the segmentation results.

In the STATISTICA 12 software environment, using the Cluster analysis module, the customer base of a telecommunications company was segmented using SOM.

The selected settings for building the model were entered in the windows of the Kohonen neural network module (Figure 4, a, Figure 4, b).

Network training begins after starting the corresponding process with the specified parameters. The dialog box displays the current training status: epoch number, network type, number of neurons, and current error value.



a) b)

Figure 4 - Kohonen neural network settings in the STATISTICA 12: determining the topological dimensions of the map (a); setting the training parameters (b).

After completing 500 full training cycles, a Kohonen map is formed, which reflects the clustering results. Each cell of the map corresponds to one of 432 neurons, and its color scale and numerical labels illustrate the number of examples matched with the corresponding node. The constructed map reflects the distribution of customers in a multidimensional feature space, which allows identifying areas with a high concentration of objects and general patterns of their placement.

The visual structure of the map allows us to assess not only the degree of fullness, but also the distribution of objects between clusters. As a result, a coordinated segmentation is formed, which is further used to detail the behavioral profiles of telecommunications company customers (Figure 5).

After completing the training of the Kohonen neural network in the STATISTICA 12 environment, the system proceeds to visualize the clustering results in the form of a two-dimensional map. Each node (neuron) of the Kohonen map corresponds to a group of customers with similar behavioral characteristics. During the training process, customers with similar activity profiles are concentrated within one or neighboring nodes, which allows identifying natural groups in the dataset without prior labeling.

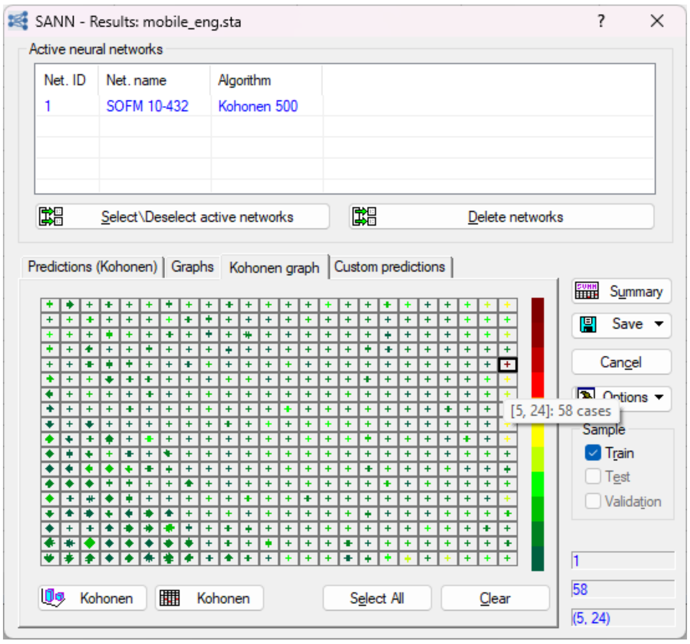


Figure 5 - Results of customer segmentation by behavioral characteristics in STATISTICA 12

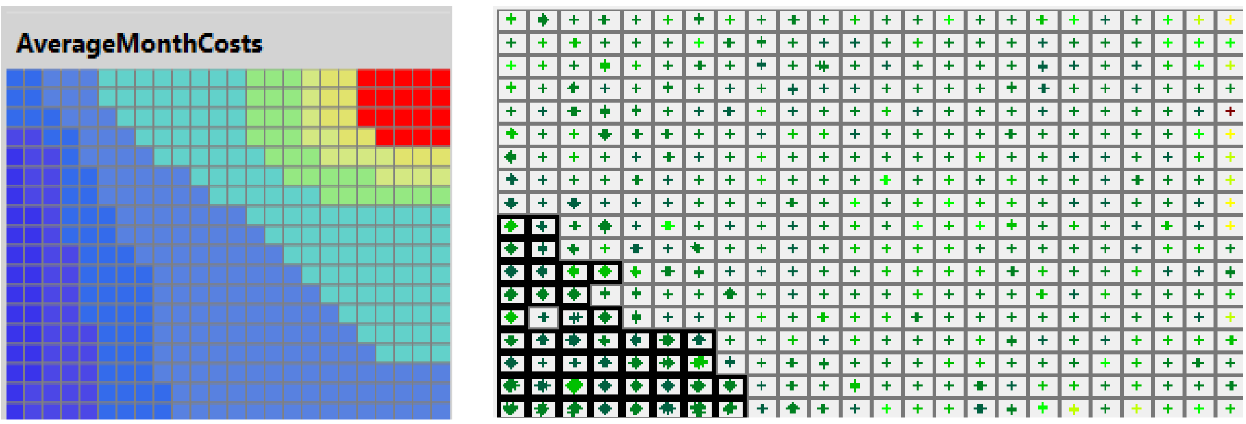
The map is presented as a rectangular grid, where the number of examples that fell into each neuron is visually marked. The density of cell filling varies, indicating the uneven distribution of clients in a multidimensional feature space. Some nodes show a high concentration of observations, indicating the presence of clearly defined segments with typical behavior. Other neurons remain empty or contain a small number of cases, which is normal for SOM models that retain their topological structure even in the case of uneven input density.

Research results

As a result of analyzing Kohonen maps obtained after segmenting the customer base of a telecommunications company using a specialized software application (Figure 3) and the Cluster analysis module of the STATISTICA 12 environment (Figure 5), four customer segments were identified, each characterized by differences in mobile communication usage profiles, in particular in terms of expenditure, call intensity, SMS activity, and other parameters.

It should be noted that in the STATISTICA environment, to obtain average values for clusters, it is necessary to additionally perform the clustering distribution procedure and perform calculations using built-in statistical functions or by exporting data for processing in external tools. In contrast, in the created application, the user can select one or more nodes on the constructed Kohonen map, after which the system automatically calculates the average values of behavioral indicators for customers matched with the selected map nodes and displays them in a special information field below the map.

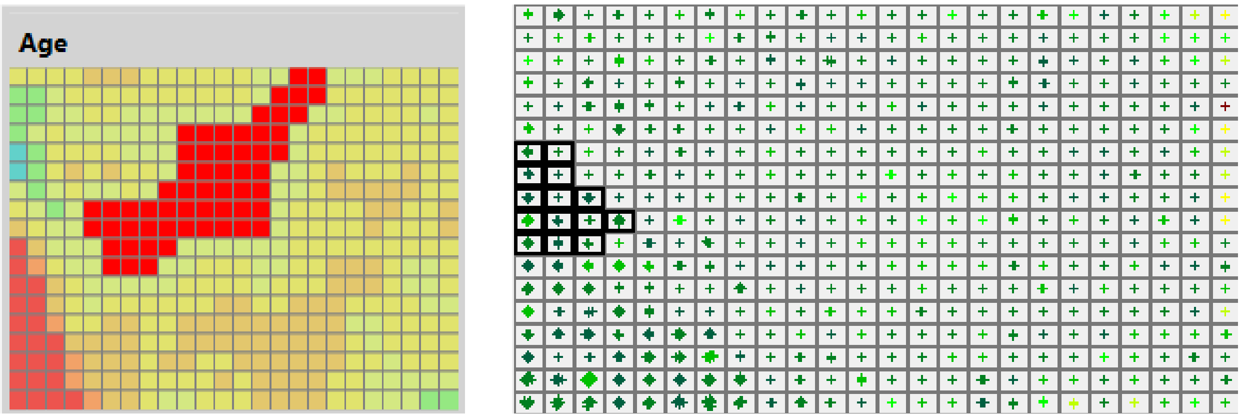
Cluster 1 combines subscribers with the highest level of spending and high voice activity. In the developed software application, the average monthly spending is 3709.85 UAH, the number of calls during the day and evening periods is 335.67 and 107.23, respectively, the average duration of calls is 12.86 minutes, and the share of calls to landlines is 25.72%. In the STATISTICA 12 system, the indicators in this segment are similar, although slightly lower: average expenses – 3206.17 UAH, number of calls during the day – 325.29, in the evening – 102.71, average call duration – 11.36 minutes, share of calls to landlines – 28.85%. In both implementations, the cluster is characterized by insignificant use of SMS (6.88 messages in the STATISTICA system and 10.15 in the developed application), which indicates a predominantly voice type of communication. Based on the set of characteristics, the segment corresponds to the category of VIP users – customers who actively use mobile communications regardless of the cost of services. On the Kohonen map, this cluster is located in the upper right corner in the developed application (Figure 6, a) and in the lower left corner in the STATISTICA system (Figure 6, b).



a) b)

Figure 6 - Cluster 1 ("VIP customers") on the Kohonen map: created software application (a); STATISTICA 12 system (b)

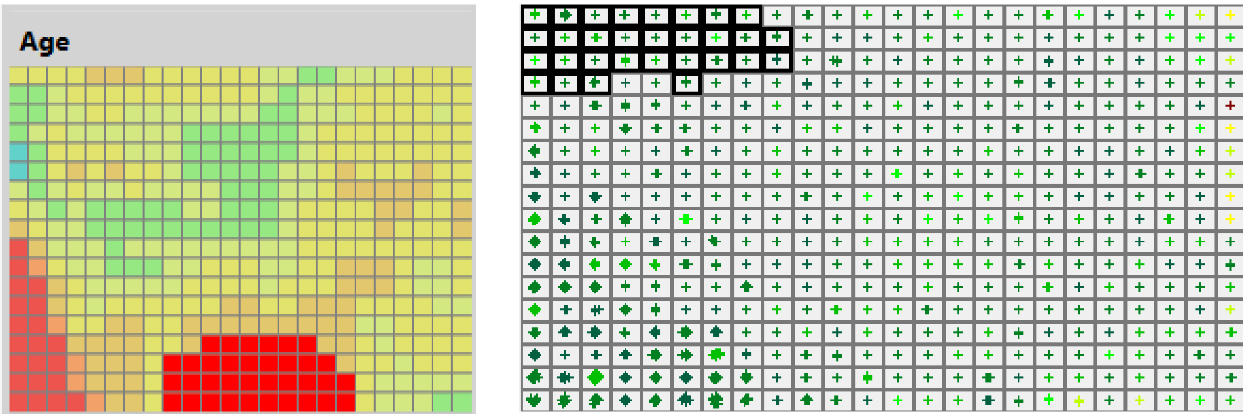
Cluster 2 represents the segment of active young people who actively use SMS. In the developed software application, this cluster is characterized by a high number of SMS messages – an average of 58.32 messages per month, a moderate level of expenditure of 928.68 UAH/month, and an average age of subscribers of 32.99 years. Similar indicators are observed in the STATISTICA 12 environment: 49.28 messages per month, 906.58 UAH in expenses, and an average age of 31.81 years. In both cases, activity is recorded during the night (23.47 in the app and 21.61 in the STATISTICA system), and the number of calls during the day is significantly lower than in the VIP segment. The average call duration is 7.65 minutes in the app and 7.33 minutes in STATISTICA. This behavior indicates a preference for text communication and a flexible style of interaction characteristic of young people. On the Kohonen map, this cluster is located in the central part of the developed application (Figure 7, a) and in the left part of the STATISTICA (Figure 7, b).



a) b)

Figure 7 - Cluster 2 ("Active Youth") on the Kohonen map: created software application (a); STATISTICA 12 system (b)

Cluster 3 brings together middle-aged subscribers with moderate mobile activity. In the created software application, the average age of users is 49.03 years, the average monthly expenditure is 496.15 UAH, and the average call duration is 5.3 minutes. In the STATISTICA 12 environment, this cluster has similar characteristics: average age – 55.45 years, spending – 291.61 – UAH/month, average call duration – 3.12 minutes. Call activity is mainly concentrated in the evening (71 calls/month). This behavior is typical for users with predictable habits focused on voice communication outside working hours. On the Kohonen map, this cluster is located at the bottom of the map in the developed application (Figure 8, a) and at the top left of the map in the STATISTICA environment (Figure 8, b).

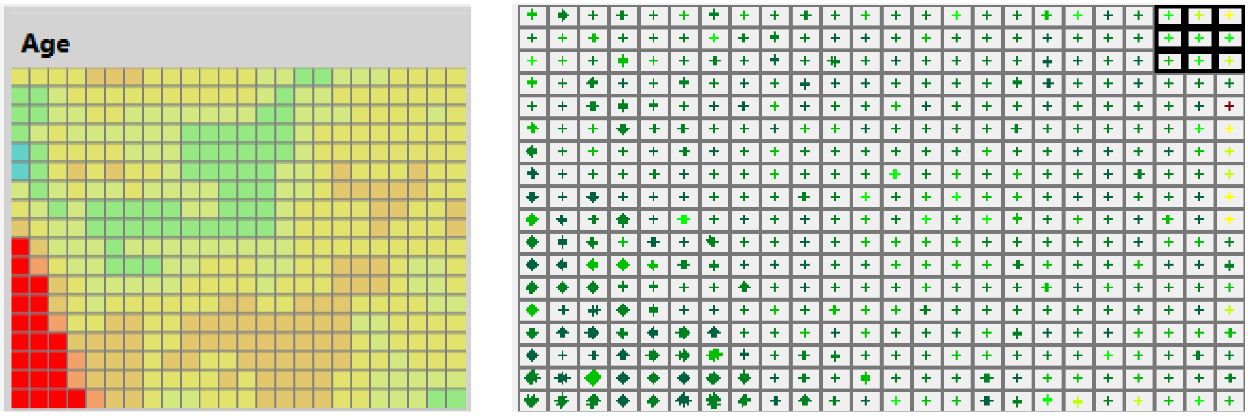


a) b)

Figure 8 - Cluster 3 ("Mature customers") on the Kohonen map: developed software application (a); STATISTICA 12 system (b)

Cluster 4 covers elderly subscribers with minimal mobile activity. In the developed application, the average age is 65.14 years, average monthly expenses are 46.79 UAH, the number of SMS is 1.48, the average duration of calls is 2.19 minutes, and the number of calls in the evening is 7.27. In the STATISTICA, this segment shows lower indicators: spending – 35.16 UAH/month, SMS – 1.52, call duration – 2 minutes, evening calls – 5.98, and the average age is 66.21 years. International and long-distance calls are nearly absent in both results. Subscribers in this cluster communicate infrequently, likely for basic needs such as short calls to relatives or emergencies. On the Kohonen map, this cluster is in the lower left corner in the developed application (Figure 9, a) and upper right in STATISTICA (Figure 9, b).

The average values of customer characteristics for each of the four clusters were determined based on the results of segmentation performed in two software environments. In the developed application, these indicators were obtained directly from the corresponding visualizations on Kohonen maps (Figure 6–9, a). In the STATISTICA 12 system, after applying the clustering propagation procedure, similar average values were calculated for each cluster (Figure 6–9, b). The generalized segmentation results from both applications are summarized in Table 1.



a) b)

Figure 9 - Cluster 4 ("Older customers") on the Kohonen map: created software application (a); STATISTICA 12 system (b)

Table 1 – Average values of customer characteristics based on data from selected cluster nodes (cl.)

Characteristic	Developed application				Developed application			
	Cl. №1	Cl. №2	Cl. №3	Cl. №4	Cl. №1	Cl. №2	Cl. №3	Cl. №4
Age (years)	44.29	32.99	49.03	65.14	45.44	31.81	55.45	66.21
Average monthly expenses (UAH)	3709.85	928.68	496.15	46.79	3206.17	906.58	291.61	35.16
Average call duration (min)	12.86	7.65	5.3	2.19	11.36	7.33	3.12	2.00
Calls during the day per month (pcs./month)	335.67	82.05	51.9	11.44	325.29	80.88	56.19	8.95
Evening calls per month (pcs./month)	107.23	85.6	71.71	7.27	102.71	99.08	71.35	5.98
Night calls per month (pcs./month)	22.16	23.47	1.39	0.11	17.81	21.61	1.63	0
Calls to other cities (number/month)	25.75	10.23	19.35	0.37	27.17	12.24	9.72	0
Calls to other countries (number/month)	1.87	0.38	0.49	0.01	2.08	0.46	0.36	0
Share of calls to landlines (%)	25.72	9.89	10.35	6.31	28.85	10.17	11.92	5.92
Number of SMS (per month)	10.15	58.32	5.54	1.48	6.88	49.28	3.07	1.52

Identifying four segments with distinct behavioral characteristics allows for a deeper understanding of the structure of a telecommunications company's customer base and the development of targeted marketing strategies tailored to the specifics of each group. Such segmentation not only reflects the structure of the customer base in terms of behavioral characteristics, but also supports management decision-making based on the analysis of this data.

Discussion of results

The results of segmenting the telecommunications company's customer base using Kohonen self-organizing maps showed a clear cluster structure that reflects the diversity of mobile communication usage patterns. Four segments were formed, each of which differs significantly in terms of expenditure, call activity, SMS usage intensity, and age characteristics. A comparison of the clusters formed in the created software application and in the STATISTICA 12 system revealed their substantive similarity: in each case, four groups with similar behavioral characteristics are clearly visible. This correspondence indicates the stability of the segmentation results when using different tools and confirms the integrity of the identified clusters.

The first cluster brings together the most active users with high spending levels. These subscribers mostly make a large number of voice calls during the day and evening, have long average call durations, and hardly ever use SMS. Their behavior is typical of the VIP segment – people who actively use mobile communications for professional needs. Given their high income and sensitivity to service quality, it is advisable to offer such customers premium rates, personalized support, and packages.

The second cluster represents the segment of younger users with moderate spending and active use of SMS. These subscribers often demonstrate nighttime activity, make fewer calls during the day, and prefer non-voice forms of communication. This behavior indicates flexibility in communication style and high adaptability to digital services. For this segment, it is advisable to develop tariff offers with favorable conditions for text messages, messenger access.

The third cluster consists of middle-aged customers with moderate voice activity. They tend to make calls in the evening, spend moderate amounts, and rarely use SMS messages. This behavior pattern indicates stable mobile communication habits focused on classic voice communication outside of working hours. This segment may be particularly sensitive to changes in tariff plans, so it is worth offering optimized packages with additional free evening minutes and stable conditions.

The fourth cluster covers the least active users, mainly elderly people. Within this segment, there are minimal expenses, short calls, very rare use of SMS, and a complete absence of long-distance or international calls. This behavior pattern indicates that mobile communications are used primarily for the basic purpose of brief contact with loved ones. For this category, it is advisable to offer simplified tariff packages with low subscription fees and a fixed amount of additional minutes.

A comparison of two clustering software tools showed that, despite certain differences in the average values of customer characteristics according to the data of the selected cluster nodes, the overall cluster structure is reproduced in both software implementations. This indicates the reproducibility of results and the stability of segments to variability in settings. The developed application implements interactive visualization of the Kohonen map with the ability to automatically calculate average values within selected clusters, which makes it convenient to perform analysis even for users without experience in working with specialized statistical tools. In STATISTICA 12, similar actions are implemented through the clustering propagation procedure with subsequent analysis using built-in tools or by exporting data to third-party tabular or analytical systems. Differences in the location of clusters on maps are explained by the initial training parameters, in particular, the random initialization of weight coefficients, which can affect the topology of the formed map.

The main advantage of the created application is the method of visualizing the Kohonen map. The method of displaying nodes on a Kohonen map in STATISTICA 12 complicates the identification of clusters, since nodes are represented as conventional graphic symbols. In contrast, the created application implements a detailed and clear graphic display of the map, which simplifies visual analysis and work with individual areas.

Thus, the clustering results obtained using Kohonen maps confirm the possibility of effectively dividing the telecommunications company's customer base into user profiles with different mobile communication usage behaviors. A comparison of the results obtained in the STATISTICA 12 system and in the created application showed a similar cluster structure, confirming the reliability of the results obtained. In particular, the developed application implements the basic functions of DSS, providing interpreted visualization of segments and simplified access to analytical information for further use in marketing strategy management.

Conclusions

1. Clustering of the telecommunications company's customer base was performed in two software environments: a specialized software application developed based on the Kohonen self-organizing map algorithm, and the Cluster analysis module in the STATISTICA 12 statistical data analysis system. In each of them, a two-dimensional Kohonen map was constructed based on the data of telecommunications company customers, which made it possible to identify the cluster structure of subscribers' distribution according to behavioral characteristics.

2. A comparison of the clustering results showed a high level of similarity in the behavioral characteristics of the selected groups. It was found that minor differences between the average values of indicators within the respective clusters indicate the consistency of the clustering results. It was shown that both types of software allowed the same customer segments to be identified: VIP users, young people, middle-aged customers, and inactive older subscribers.

3. A comparative analysis of the functional capabilities of the created specialized application and the Cluster analysis module in the STATISTICA 12 system was carried out, taking into account the practical requirements for processing customer data in the telecommunications industry. The created application implements functionality that allows it to be used as an element of a decision support system: in particular, it provides automatic calculation of average values for selected nodes of the Kohonen map, interactive interaction with clustering results, and convenient visual presentation of data. This makes the created software tool suitable for quickly forming conclusions without involving external processing tools. In turn, the use of the STATISTICA 12

system involves a more complex data processing procedure, limited visual informativeness of the Kohonen map, and the need for additional actions to obtain descriptive statistics, which reduces the convenience of using the tool in applied tasks of customer base analysis.

References

1. Customer Segmentation in the Telecom Industry. URL: https://www.subex.com/glossary/customer-segmentation-in-the-telecom-industry/

2. Shumeiko, O. O., & Sotnik, S. L. (2012). Intelligent Data Analysis (Introduction to Data Mining): Textbook. Dnipropetrovsk: Belaya E. A. – 212 pp.

3. Analysis of the Customer Base: Identifying Potential Opportunities. URL: https://blog.lp-crm.biz/uk/crm-system/analiz-klientskoj-bazy-nahozhdenie-potenczialnyh-vozmozhnostej/

4. Kotler, F., Keller, K. L., Pavlenko, A. F., et al. (2008). Marketing Management: Textbook. First Ukrainian adapted edition. Kyiv: Himdjest – 720 pp.

5. Power, D. J. (2002). Decision Support Systems: Concepts and Resources for Managers. Quorum Books.

6. Wedel, M., & Kamakura, W. A. (2000). Market Segmentation: Conceptual and Methodological Foundations (2nd ed.). Boston: Kluwer Academic Publishers.–382 pp.

7. What is Market Segmentation and How to Use It in Marketing. URL: https://sendpulse.ua/blog/what-is-market-segmentation/

8. Jain, A. K. (2010). Data clustering: 50 years beyond k-means. Pattern Recognition Letters, Vol. 31, No. 8, pp. 651–666. DOI: 10.1016/j.patrec.2009.09.011

9. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Multivariate Data Analysis (7th ed.). Pearson Education. – 785 pp.

10. Oliynyk, S. V., & Boyko, B. M. (2023). Comparison of clustering methods for creating target customer groups in a dataset. Scientific Bulletin of NLTU of Ukraine, Vol. 33, No. 5, pp. 77–83. DOI: 10.36930/40330510.

11. Hand, D., Mannila, H., & Smyth, P. (2001). Principles of Data Mining. Cambridge, Massachusetts: MIT Press. – 546 pp.

12. Farahani, F. J., & Tabibian, S. (2024). Innovative customer segmentation based on multi-step sequential deep clustering in the telecommunication industry. Proceedings of the 14th International Conference on Computer and Knowledge Engineering (ICCKE), pp. 1–6. DOI: 10.1109/ICCKE65377.2024.10874819.

13. Twomey, P., & Cadman, R. (2002). Agent based modeling of customer behavior in the telecoms and media markets. Info, Vol. 4, No. 1, pp. 56–63. DOI: 10.1108/14636690210426640.

14. Sarkar, M., Roy, P. A., & Chowdhury, F. R. (2024). Optimizing marketing strategies with RFM method and k-means clustering-based AI customer segmentation analysis. *Journal of Business and Management Studies*, Vol. 6, No. 2, pp. 54–60. DOI: 10.32996/jbms.2024.6.2.5.

15. Gomes, M. A., & Meisen, T. (2023). A review on customer segmentation methods for personalized customer targeting in e-commerce use cases. Information Systems and e-Business Management, Vol. 21, pp. 527–570. DOI: 10.1007/s10257-023-00640-4.

16. Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. Biological Cybernetics, Vol. 43, No. 1, pp. 59–69. DOI: 10.1007/BF00337288.

17. Schreck, T., Bernard, J., von Landesberger, T., & Kohlhammer, J. (2009). Visual cluster analysis of trajectory data with interactive Kohonen maps. Information Visualization, Vol. 8, No. 1, pp. 14–29. DOI: 10.1057/ivs.2008.29.

18. Ultsch, A., & Vetter, C. (1994). Self-organizing feature maps versus statistical clustering methods: A benchmark. Philipps University of Marburg, p. 15.

19. Wehrens, R., & Buydens, L. (2007). Self- and super-organizing maps in R: The kohonen package. *Journal of Statistical Software*, Vol. 21, No. 5, p. 19.

20. StatSoft, Inc. Electronic Statistics Textbook. URL: https://www.freetechbooks.com/electronic-statistics-textbook-t932.html

21. Self-Organizing Maps Explained. URL: https://medium.com/%40becaye-balde/self-organizing-maps-explained-9bb88b9a620a/

Article sent: 13.09.2025