Modeling of the Patient's Cognitive Status Based on Fuzzy Clustering of Psychometric and Neurobiological Data

Tatiana Avdeenko [a], Ekaterina Kirianova [a], Olga Razumnikova [a], Irina Tarasova [b], Olga Trubnikova [b]

[a] *Novosibirsk State Technical University, 20, Karla Marksa ave., Novosibirsk, 630073, Russia*
[b] *Research Institute of Complex Problems of Cardiovascular Diseases, 6, Sosnoviy blvd, Kemerovo, 650002, Russia*

Abstract

In present paper we consider the problem of mathematical processing of medical data and investigate the performance of clustering algorithms on psychometric and neurophysiological data of patients who underwent coronary artery bypass grafting. Cluster analysis of patients can help in determining patterns of deterioration of cognitive functions and understanding in what way the age, education and indicators of cognitive status (MMSE and CCI) are related to electroencephalogram (EEG) results. Here we compare between traditional clustering and fuzzy clustering algorithms. Fuzzy clustering of a dataset is done using the fuzzy C-means (FCM) method. Algorithms for using mixed methods of clustering and analysis of variance to solve the problem of analyzing risk factors for cognitive pathologies are considered. The results of the algorithms performance for clustering the adaptive cognitive potential of the functional activity of the brain are presented.

Keywords [1]

cluster analysis, fuzzy clustering method, mixed clustering methods, analysis of variance

1. Introduction

In the modern world we observe permanent growth in the volume of heterogeneous data. In this regard the problems arise with their processing and research. The problem of mathematical processing of medical data is not an exception. The latter requires choice and permanent improvement of methods and algorithms for analysis of data being specific in this area. Compared to well-developed clinical field of medical research, the area of statistical inference needs more attention. The use of analysis algorithms, especially those associated with new intelligent technologies, makes it possible to reveal informative relationships, features, as well as new knowledge about the objects that is not obvious to an expert. The latter can significantly increase the likelihood of correct and quick diagnosis of diseases and complications of patients.

One of the modern algorithms for data analysis is cluster analysis - a section of unsupervised machine learning. It is a powerful tool that allows you to identify the main features, hidden patterns and relationships in a variety of heterogeneous information, which is useful for analyzing biometric or neurophysiological data, which often include nominal factors. One of peculiarities of cluster analysis is the lack of a priori information about the structure of groups (clusters) and their number. This means that there simply does not exist an unambiguously correct and high-quality method to split a dataset. Also, almost all known methods essentially depend on the metric used, the choice of which is

YRID-2020: International Workshop on Data Mining and Knowledge Engineering, October 15-16, 2020, Stavropol, Russia
EMAIL: avdeenko@corp.nstu.ru (Tatiana Avdeenko); katmc@yandex.ru (Ekaterina Kirianova); razoum@mail.ru (Olga Razumnikova); iriz78@mail.ru(Irina Tarasova); olgalet17@mail.ru(Olga Trubnikova)
ORCID: 0000-0002-8614-5934 (Tatiana Avdeenko); 0000-0001-9018-8964 (Ekaterina Kirianova); 0000-0002-7831-9404 (Olga Razumnikova); 0000-0002-6391-0170 (Irina Tarasova); 0000-0001-8260-8033 (Olga Trubnikova)

usually determined by an expert.

This article explores the possibilities and compares existing cluster analysis algorithms in relation to psychometric and neurobiological data on patients undergoing coronary bypass surgery. The observations contain information on 39 right-handed men, aged from 52 to 60 years, diagnosed with coronary heart disease (CHD) [1], for whom multichannel computed electroencephalography (EEG) was performed, quantitatively assessing changes in the cerebral cortex. EEG results were obtained 3-5 days before coronary artery bypass grafting (CABG) surgery, as well as 7-10 days after the surgery. To divide the subjects of the sample into groups, we used the EEG values with eyes closed in theta-1 (4-6 Hz), alpha-1 (8-10 Hz) and beta-2 (20-30 Hz) bands, as well as age indicators, education, the sum of points on a Mini Mental State Examination (MMSE) and a complex indicator of cognitive status (CCI), which is a summary characteristic of voluntary attention, short-term memory and executive functions [1].

EEG results are considered as a possible risk factor for cognitive pathologies, since there is an assumption about their relationships with CHD. The initial manifestations of such abnormalities are often not diagnosed in clinical practice, although their timely treatment can help prevent dementia, the most severe form of mental impairment [2]. The literature contains extensive information on the causes and risk factors for their development, but they are extremely contradictory and do not give a clear answer to the question of their relationships with the EEG values. Cluster analysis of patients can help in determining the patterns of cognitive decline and will help to understand how the factors of age, education and MMSE are related to EEG results.

Formally, the clustering problem can be described as follows: there is a data set $X = \{x_1,...,x_n\}$ and a selected similarity metric $\rho = \rho(x, y)$, which determines the numerical value of the distance between two objects *x* and *y* based on their properties and parameters. It is necessary to split the sample into groups (clusters) so that each of them consists of objects that are close in terms of the metric criterion, and the groups themselves are as different as possible. Then, by the clustering algorithm we mean a function $a : X \to Y$, which assigns the cluster identifier $y \in Y$ to all objects in the sample. In this paper we investigated the possibility of dividing patients into 3 clusters. However, fuzzy analysis shows that further research is of interest, implying a division into 2 clusters, on the basis of which additional interesting conclusions can be drawn in the future.

2. Methods of cluster analysis

2.1. Traditional methods

Cluster analysis was originated in anthropology by Driver and Kroeber in 1932 and introduced to psychology by Zubin in 1938 [3]. And also, it was used by Cattell beginning of 1943 [4] for trait classification in personality psychology [5].

Traditional clustering algorithms break up a dataset into many disjoint groups. Any such algorithm must satisfy two requirements. First, each cluster must contain at least one observation. Second, each observation must belong to one and only one cluster.

The main idea of traditional methods is to use an iterative regrouping technique, given an initial separation, which attempts to improve the separation by moving objects from one group to another. Under the general concept of «improvement» we mean: objects of one cluster must be close or related to each other; objects of different clusters must be far from each other, or have different meanings.

There are several features inherent in all traditional clustering algorithms:

- can detect only spherical clusters,
- use different distance metrics to determine cluster membership,
- to represent a cluster, a certain object (centroid, medoid) is needed,
- effective only for small datasets.

Algorithm K-Means

Let a dataset *X* contain *N* objects in Euclidean space. The algorithm distributes objects into *k* clusters $C_1,...,C_k$, that is, $C_i \subset X$ and $C_i \cap C_j = \varnothing$ for $1 \le i, j \le k$. For each cluster, its center is determined being an object called the centroid. The method is aimed at minimizing the objective function, which is the total square deviation of cluster objects from their centroids:

$$E = \sum_{i=1}^{k} \sum_{p \in C_i} ||| x - \mu_i ||^2 \quad (1)$$

,

where μ is the centroid of all vectors *x* from the cluster C_i.

The result of applying K-Means is a local minimum. This means that several runs of the algorithm with random initial centers can give potentially better results. Modifications of the method can be used to obtain more accurate partitions.

Algorithm K-Means++

The disadvantage of the K-Means algorithm is the sensitivity to initialization of the initial centroids. If the centroid is initialized as an object remote from the main data collection, then it may simply not have neighbors associated with it. Likewise, more than one centroid can be initialized in the same cluster, resulting in poor partitioning.

This problem can be solved by changing the initialization phase in KMeans ++. This modification tends to place the initial centroids as far apart as possible, which increases the chances of initially detecting centroids that are in different clusters [6]. At the initialization stage, the probability of choosing each next centroid is proportional to its distance to the nearest previously selected centroid, which contributes to a significant decrease in the error of the final clustering result. After the initialization phase, the main standard K-Means iterative process: Although the modified initial centroid selection takes additional time, K-Means ++ has fast convergence and improved stability.

Algorithm partitioning around medoids (PAM)

In this method, each cluster is represented by one of the data objects called medoids, which is similar to the centroid concept in K-Means. The difference is similar to the difference between mean and median: where mean indicates the average of all collected data items, and median indicates the value around which the objects are uniformly distributed. Also, PAM is more robust to noise than K-Means algorithm.

The original PAM algorithm randomly selects medoids from sample objects. To increase the sustainability of the results, in this work we change this stage replacing standard initialization to the initialization stage of the K-Means ++ algorithm.

The main difference between PAM and K-Means or K-Means ++ is in the selection of new medoids, which are always objects of the data set. The main disadvantage of the algorithm is its computational complexity and low quality of work on large amounts of data.

2.2. Fuzzy clustering method

Fuzzy logic is a generalization of classical logic when the truth value takes not only two values "true" and "false", but a continuous set of values from the interval [0,1]. Thus, fuzzy logic is a form of flexible, inaccurate computation that mimics human decision making [7-9].

Fuzzy logic is used in data analysis in both supervised and unsupervised learning. In supervised learning, it is used, for example, in the case of a fuzzy modification of the decision tree method [10], as well as in the automatic formation of the knowledge base in the form of fuzzy rules from the case base [11]. Fuzzy clustering methods are also related to the unsupervised machine learning.

The main peculiarity of fuzzy clustering is that it allows an object to belong to each cluster with a certain degree of membership ranging from 0 to 1. The most common methods of fuzzy clustering are aimed at minimizing the objective function, the main parameters of which are degrees of membership and parameters that determine localization and shape of the clusters. One of such methods is the Fuzzy C-Means (fuzzy means) algorithm [12], which minimizes the objective function

$$H_{FCM} = \sum_{k=1}^{c} \sum_{i=1}^{n} \mu_{ik}^{m} dist(x_i, v_k)^2, \quad (2)$$

where $dist(x_i, v_k)$ is a selected metric of the distance between the object x_i and the cluster center v_k, μ_{ik} – the degree of belonging the object i to the cluster k satisfying

$$\sum_{k=1}^{c} \mu_{ik} = 1, \tag{3}$$

m is an exponential weight, or "fuzziness index", which controls the blurring of clusters. The value m is usually set in the range from 1 to 2. Such a choice can be considered as a trade-off between assuming the degree of fuzziness in the dataset and avoiding the tedious calculation of its value. However, if this index is carefully adjusted, the algorithm can be optimized to account for the characteristic noise present in the dataset.

Minimization of the objective function H_{FCM} is performed by the method of alternating optimization. The cluster membership of each object is updated using the following formulas:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ki}}{d_{ji}} \right)^{\frac{2}{m-1}}}, \tag{4}$$

where $k \in [1, c]$, $i \in [1, n]$. The cluster center is recalculated as:

$$v_k = \frac{\sum_{i=1}^{n} (\mu_{ik}^m x_i)}{\sum_{i=1}^{n} \mu_{ik}^m}. \tag{5}$$

The algorithm can be described as follows:

1. The elements of the membership matrix are randomly generated from the degrees of membership μ_{ik}, satisfying the relation (3);
2. The centers of the clusters are calculated by the formula (5);
3. New elements of the membership matrix are calculated by the formula (4);
4. Steps 2-3 are repeated until the convergence of the algorithm, which is determined by checking the relation:

$$\| U^{(k+1)} - U^{(k)} \| < \varepsilon, \tag{6}$$

where *U* is the membership matrix, ε is the specified accuracy such that $\varepsilon > 0$.

2.3. Mixed clustering algorithms

The above clustering methods are used primarily for quantitative features. To get the results on a full multidimensional mixed valued data set, modified methods are needed, which will allow working with both nominal and ordinal features. One of these methods is K-Prototypes related to non-fuzzy (crisp) clustering [13]. It is based on a combination of the K-Modes and K-Means algorithms, repeating the idea of iterative center refinement, and removing the ordinal factor constraints. As it is also indicated in [14] that, in the general case, K-Prototypes works faster than standard K-Means, since fewer iterations are required for its convergence.

The peculiarity of K-Prototypes consists in a new measure of similarity for nominal features and a frequency method for updating cluster centers to minimize the cost function expressed by the formula

$$E = \sum_{i=1}^{n} \sum_{j=1}^{k} u_{ij} d(x_i, \mu_j), \tag{7}$$

where $x_i, i = 1,...,n$ are observations in the sample, $\mu_j, j = 1,...,k$ are cluster centers called prototypes, u_{ij} are elements of binary partition matrix $U_{n \times k}$ satisfying the expression

$$\sum_{j=1}^{k} u_{ij} = 1, \forall i . \tag{8}$$

The similarity measure function has the form

$$d(x_i, \mu_j) = \sum_{m=1}^{q} (x_i^m - \mu_j^m)^2 + \lambda \sum_{m=q+1}^{p} \delta(x_i^m, \mu_j^m), \tag{9}$$

where m is an index over all variables in the dataset, with the first q variables being ordinal, and the rest $p-q$ are nominal. Moreover, $\delta(a,b)=0$ for $a=b$, $\delta(a,b)=1$ for $a \neq b$. Measure (9) corresponds to the weighted sum of the Euclidean distance between two points in metric space for numerical variables and the simple correspondence distance (number of mismatches) for nominal variables, λ is a predetermined weight of nominal variables. When, $\lambda = 0$ the influence of nominal variables disappears and only ordinal variables are taken into account, as in the traditional K-Means algorithm.

3. Results of cluster analysis of psychometric and neurophysiological data

3.1. Comparison of traditional methods

In the data set under study, there are two types of traits: nominal, to which the education factor belongs, and ordinal, which include age, biometric and cognitive parameters. Since all traditional methods described above show themselves best on samples with ordinal characteristics, the education factor will not be taken into account for them.

To evaluate the results of traditional methods of cluster analysis, an optimality criterion was introduced, by which we mean the value of the sum of the squares of the distances of each data object to the center of the cluster:

$$E = \sum_{j=1}^{k} \sum_{i=1}^{n} \| x_i - \mu_j \|^2, \tag{10}$$

where μ is the center of all vectors x from the cluster j.

According to the results from Tables 1, 2, and 3, it can be seen that the considered traditional methods uniformly distribute the objects of the dataset, giving preference to the first and the third clusters.

Table 1

Results of the K-Means algorithm

Test number		1	2	3	4	5	6	7	8	9	10
Number of patients in each cluster	1	20	23	22	18	19	16	23	19	21	22
	2	9	6	6	6	6	6	5	6	9	5
	3	10	10	11	15	14	17	11	14	9	12
Number of iterations		5	6	4	4	7	2	4	7	3	4
Optimality criterion		241.5	234.4	238.9	235.7	236.5	236.3	240.1	235.0	259.6	238.6

Table 2

Results of the K-Means++ algorithm

Test number		1	2	3	4	5	6	7	8	9	10
Number of patients in each cluster	1	20	22	26	18	23	21	17	25	26	17
	2	9	6	3	6	6	6	6	3	1	10
	3	10	11	10	15	10	12	16	11	12	12
Number of iterations		5	6	1	2	4	1	4	3	2	4
Optimality criterion		241.5	238.9	250.3	235.1	234.4	236.1	234.8	251.8	296.6	257.9

Table 3
Results of the PAM algorithm

Test number		1	2	3	4	5	6	7	8	9	10
Number of patients in each cluster	1	18	19	22	20	22	21	19	18	23	15
	2	6	6	1	7	6	6	4	9	5	12
	3	15	14	16	12	11	12	16	12	11	12
Number of iterations		1	2	1	1	2	1	1	2	1	1
Optimality criterion		291.9	331.9	345.4	338.6	296.6	290.1	348.7	328.9	292.2	330.0

Almost every time any of the three algorithms was run, the second cluster included an average of 6 to 9 objects of the whole set, which can be explained by their better stability in comparison with the objects of the first and third clusters. By the concept of stability, we mean how different the resulting partitions into the clusters are as a result of multiple applications of algorithms for the same data. Confirmation of the conclusions about stability can be obtained by analyzing the results of fuzzy clustering, which allows you to clearly see which objects are prone to moving from one cluster to another.

3.2. Fuzzy clustering results

The results of the FCM algorithm were obtained in the form of a membership matrix of the studied dataset, for which the accuracy parameter $\varepsilon = 0.001$ in relation (6) and the fuzziness coefficient $m = 1.5$ were used. The criterion value equal to 0.000949 was achieved with 35 iterations. The majority of the sample consisted of 24 observations proved to have a degree of belonging to one cluster above 0.8. A number of objects have almost equal degree of belonging to both the first and third clusters (the difference is no more than 0.1). Such results can be explained by the semantic content of the fuzziness coefficient, which is responsible for the blurring of clusters. As its value increases, the area covered by the cluster also increases, which means more intersections with other clusters, and less belonging to the same cluster for objects close to the edges of the area.

In Figure 1, the ratios of the membership matrix values for each patient in the studied dataset are displayed as a normalized histogram for clarity.

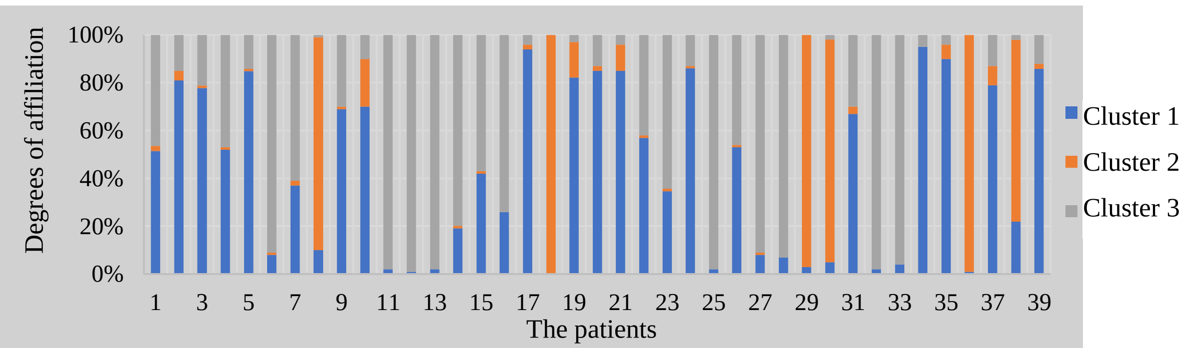

Figure 1: The membership degrees of observations of the complete data set at m=1.5

To analyze the distribution of the number of observations, each object was assigned a cluster label, determined by the maximum degree of membership function. The most numerous was the first cluster, in which there were 19 observations. The third includes 14, and the second only 6 observations.

Now consider the results with a fuzzy coefficient $m = 2$ and a criterion value equal to 0.000895, achieved at 26 iterations. Figure 2 shows a normalized histogram of the ratio of the membership matrix values for each patient in the studied dataset with m = 2.

After increasing fuzzy coefficient, the number of observations with a degree of membership higher than 0.8 decreased almost four times, to 7. The number of observations with equal membership increased to 6. The distribution of objects by clusters turned out to be almost the same as with coefficient $m=1.5$: the first cluster was 20 observations, the second was 6, the third was 13. At the same time, only one patient with number 15 made the transition from the third to the first cluster, and the patients of the second cluster remained unchanged. These results confirm the assumption of good stability of the second cluster, which means a greater similarity between its objects.

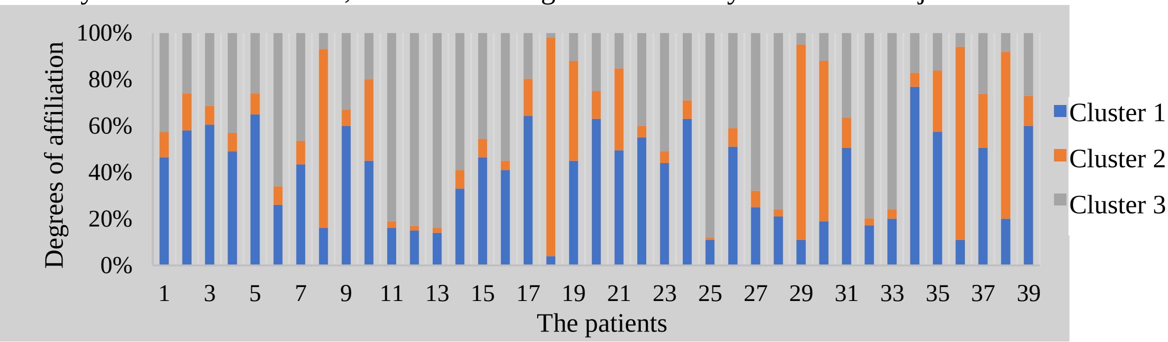

Figure 2: The membership degrees of observations of the complete data set at m=2

There are two implementations of K-Prototype. The first was developed by Z. Huang in 1997 [14], and the second one by F. Cao in 2009 [15]. The differences are only in the way prototypes are initialized, similar to the difference between K-Means and K-Means ++. The selection of centers in the Huang method is random, and the Cao method is based on maximizing the density to obtain the most accurate initial prototype.

In this work, the Cao method was used, the results of which are displayed in Table 4.

Table 4
Cluster characteristics

Before surgery					After surgery				
№	**N**	**Mean age**	**Mean MMSE**	**Mean CCI**	**№**	**N**	**Mean age**	**Mean MMSE**	**Mean CCI**
					Theta-1 band				
1	17	56,529	27,294	0,569	1	14 (5)	56,714	27,642	0,566
2	6	56,833	28,333	0,538	2	6 (1)	57,166	28	0,502
3	16	57,125	27,687	0,438	3	19 (3)	56,789	27,473	0,473
					Alpha -1 band				
1	15	56,8	26,933	0,509	1	14 (7)	57,571	27,785	0,494
2	11	56,909	28,272	0,538	2	14 (2)	56,5	27,142	0,502
3	13	56,769	27,846	0,489	3	11 (5)	56,272	28	0,542
					Beta -2 band				
1	20	55,45	27,15	0,504	1	16 (11)	56,875	26,875	0,503
2	9	58,111	28,222	0,499	2	12 (1)	57,090	28,636	0,508
3	10	58,4	28	0,534	3	11 (7)	56,5	27,666	0,523

It provides information on each band before and after coronary artery bypass grafting (CABG) surgery, namely: the number of patients (N) in the cluster, the average age of the selected group **Mean age**, the average **Mean MMSE** and **Mean CCI**. The number of patients who moved to another cluster after surgery is indicated in parentheses.

The stability of observations of the second cluster is noticeable, which showed the highest value of MMSE before the operation. The largest number of observations made the transition to other clusters at Beta-2 band. There was also a decrease in the level of CCI in the Alpha-1 band after surgery between the patients with secondary education, while patients with higher education showed a less pronounced decline in cognitive status. This may mean that people that are more educated have greater neuronal reserves, which gives them the opportunity to compensate for the effects of coronary heart disease.

3.3. ANOVA of the fuzzy clustering results

In this research fuzzy clustering of the dataset into three groups was carried out using the FCM method [16]. Of 120 patients, only nine were in the same cluster for all three bands. Of these nine, two patients belong to the third cluster and seven belong to the first one. 24 patients were assigned to different clusters for each band. All others belonged to the same cluster in two out of three bands. Figures 3, 4 and 5 show the membership values for each band.

All three graphs show the band of corresponding cluster, the degrees of membership of which form a curve of the hyperbolic type, that is, more than half have a very low degree of membership, and the rest - high, which is clearly seen in Figures 3 and 4. This means that each cluster in a certain band has a more pronounced similarity of observations.

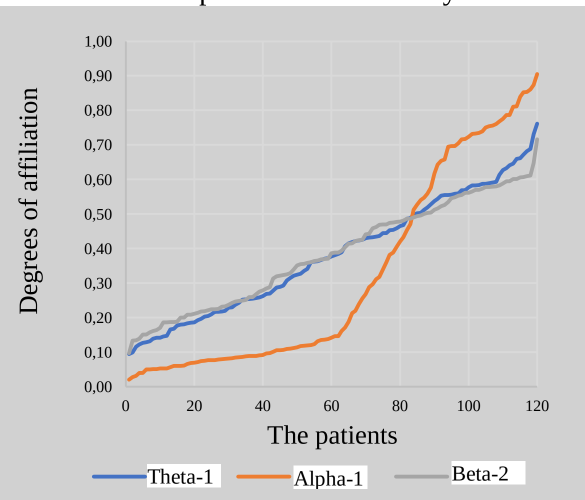

Figure 3: Cluster 1 membership degrees

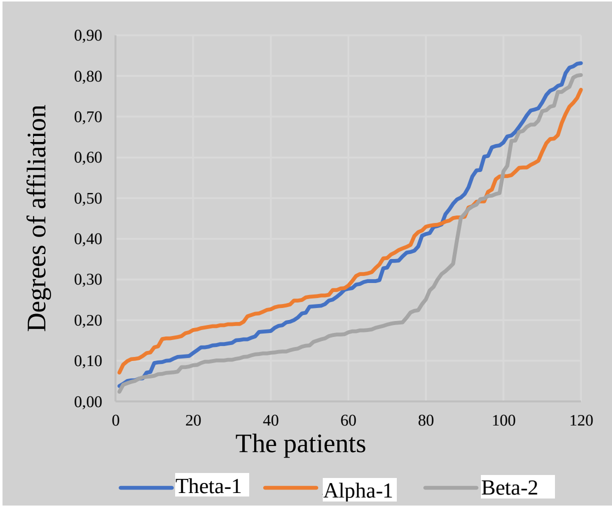

Figure 4: Cluster 2 membership degrees

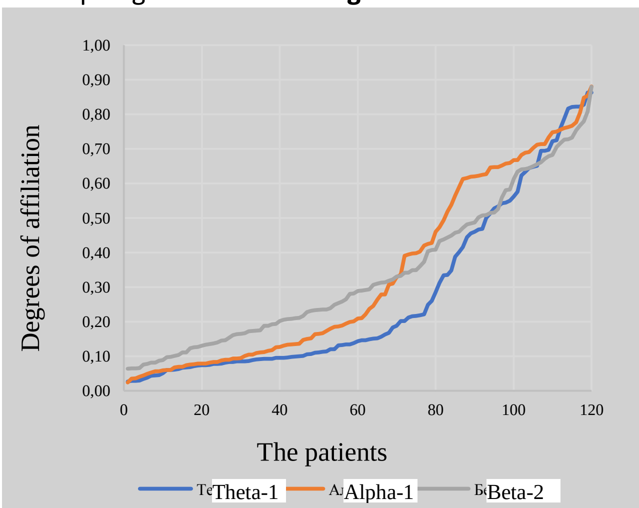

Figure 5: Cluster 3 membership degrees

To assess the stability of observations at different fuzzy coefficients of the FCM algorithm, a comparative analysis was carried out, during which similar changes in the quantitative ratio in the clusters were revealed. After reducing the fuzzy coefficient to 1.5, the number of observations with a degree of belonging above 0.8 increased almost six times, to 72, which constituted the majority of the sample. The distribution of objects by clusters changed: the first cluster consisted of 59 observations, the second - 33, the third - 28. At the same time, the patients of the second cluster remained unchanged, as well as in the sample of 60 observations. These results confirm the assumption of good stability of the second cluster, which means a greater similarity between its objects.

To determine the significance of differences in the mean values of the resulting partition, one-way analysis of variance (ANOVA) was applied for each indicator, relative to all frequency bands (Table 5).

Table 5
ANOVA Results for Three Patient Groups

Variable	F	p	Cluster			F	p	Cluster		
			1	2	3			1	2	3
Theta-1 band						**Beta-2 band**				
Age	5,36	**0,00**	57,84	<u>54,82</u>	58,70	8,54	**0,00**	58,71	55,13	59,25
Education	2,76	**0,06**	1,22	1,18	1,41	2,40	**0,09**	1,14	1,28	1,37
MMSE	0,98	0,37	27,32	27,15	27,62	1,74	0,18	27,57	27,09	27,59
CCI	0,55	0,58	0,49	0,53	0,49	0,36	0,70	0,51	0,50	0,48
Power	203,22	0,00	0,24	<u>0,51</u>	0,04	250,74	0,00	-0,64	-0,44	-0,18
Alpha-1 band										
Age	0,33	0,71	57,78	57,22	56,79					
Education	0,55	0,58	1,29	1,30	1,21					
MMSE	0,26	0,77	27,41	27,45	27,23					
CCI	0,33	0,72	<u>0,48</u>	0,51	0,51					
Power	304,21	0,00	1,38	0,21	0,83					

The table 5 highlights statistically significant differences between the age factor and the EEG parameters (parameter p), and also highlights important parameters of these factors for clusters.

It was found that clustering showed a correspondence with the greater power of the theta-1 band in younger individuals (cluster 2 for theta-1 in Table 5), which means an age-related decrease in the power of low-frequency EEG band. Also, a lower CCI value was recorded at the alpha-1 band (cluster 1 for alpha-1 in Table 5), which gives an idea of cognitive deficit.

An increase in the power of the beta band and a decrease in alpha activity may be confirmation of the fact of violations of regional neural interactions. There is an assumption that for patients with low cognitive status there is a rearrangement of electrical brain activity due to prolonged coronary heart disease. However, further studies are required for cluster analysis of different EEG bands and additional study of the influence of other factors.

4. Acknowledgements

The research is supported by Ministry of Science and Higher Education of Russian Federation (project No. FSUN-2020-0009).

5. References

[1] I.V. Tarasova, O.A. Trubnikova, I.N. Kukhareva, O.L. Barbarash, L.S. Barbarash, Annual dynamics of neurophysiological indicators in patients undergoing coronary artery bypass grafting with artificial circulation // Complex problems of cardiovascular diseases. – 2015, no.1, p. 18-24.

[2] Yu.A. Mezentsev, O.M. Razumnikova, I.V. Tarasova, O.A. Trubnikova, On some problems of clustering big data by minimax and additive criteria, application in medicine and neurophysiology // Information technologies: scientific and technical. zhurn. – 2019, T. 25, no.10, p. 602-608.

[3] J. Zubin, A technique for measuring like-mindedness // The Journal of Abnormal and Social Psychology. – 1938, vol. 33, p. 508–516.
[4] R.B. Cattell, The description of personality: bic traits resolved into clusters // Journal of Abnormal and Social Psychology. - 1943, vol. 38, no.4, p. 476–506.
[5] X. Chen, Y. Zhou, Q. Luo, A Hybrid Monkey Search Algorithm for Clustering Analysis // The Scientific World Journal Volume – 2014 – Article ID 938239, 16 p.
[6] B. Aubaidan, M. Mohd, M. Albared, Comparative study of k-means and k-means++ clustering algorithms on crime domain // Journal of Computer Science. – 2014, no.10, p. 1197–1206.
[7] T.J. Ross, Fuzzy logic with engineering applications, Third Edition. – John Wiley and Sons, 2010. – 606 p.
[8] D. Dubois, H. Prade, Fuzzy sets and systems theory and applications. – London: Academic Press, 1980. – 411 p.
[9] L.A. Zadeh, Fuzzy Sets // Information and Control. – 1965, vol.8, p.338-353.
[10] S. B. Begenova, T.V. Avdeenko, Building of fuzzy decision trees using ID3 algorithm // Journal of Physics: Conference Series. - 2018. – Vol. 1015: International conference information technologies in business and industry, Tomsk, 18–20 Jan. 2018. – Art. 022002 (7 p.). - DOI: 10.1088/1742-6596/1015/2/022002.
[11] T.V. Avdeenko, Automatic knowledge acquisition in the form of fuzzy rules from cases for solving classification problem // Fuzzy expert systems and applications in agricultural diagnosis. - [Hershey]: IGI GLOBAL, 2019, p. 33-49. - DOI: 10.4018/978-1-5225-9175-7.ch003.
[12] A. Gosain, S. Dahiya, Performance analysis of various fuzzy clustering algorithms: a review // Procedia Computer Science. – 2016, no.79, p. 100–111.
[13] H. Park, J. Lee, C. Jun, A K-means-like algorithm for K-medoids clustering and its performance // Department of Industrial and Management Engineering, POSTECH. – 2006, no.31, p. 790–784.
[14] Z. Huang, Clustering large data sets with mixed numeric and categorical values // CSIRO Mathematical and Information Sciences. – 1997, p. 21–34.
[15] F. Cao, J. Liang, L. Bai, A new initialization method for categorical data clustering // Expert Systems with Applications: An International Journal. – 2009, no.36, p. 10223–10228.
[16] U. Kaymak, Extended fuzzy clustering algorithms / U. Kaymak, M. Setnes // Erim Report Series Research in Management, November 2000, 24 p.