# Using Cluster Analysis for Author Classification of Albanian Texts: A Study on the Effectiveness of Stop Words

#### DENISA KAÇORRI , ALBINA BASHOLLI , LUELA PRIFTI Department of Mathematical Engineering, <sup>1</sup>Polytechnic University of Tirana, Faculty of Mathematical Engineering and Physics Engineering ALBANIA

*Abstract:* - Cluster analysis is a statistical approach that identifies uniform clusters within data. The closeness of data is measured quantitatively using distance functions. Specifically for text data mining, clustering serves as a method of categorization of words based on the similarity of their occurrence within texts and classifying texts by topics or author. Hierarchical clustering is a powerful technique for identifying natural groupings within datasets, which can be especially useful for unsupervised text classification. This paper aims to utilize cluster analysis to establish Albanian texts clusters by authors. Using agglomerative hierarchical clustering we classify Albanian texts by authors according to the similarity of their word frequency. The similarity of texts is evaluated using cosine and Euclidean distances. Considering two study cases, respectively with and without Albanian stop words we conclude that the best clustering by authors of the Albanian documents is achieved with 87% accuracy using Ward's method with cosine distance in the case of study by removing stop words.

Key-Words: - Clustering, text classification, Albanian text, stop words.

Received: May 5, 2022. Revised: August 9, 2023. Accepted: September 11, 2023. Available online: October 19, 2023.

## **1** Introduction

Clustering is a statistical technique that identifies consistent sets within information and is applied in various areas. Cluster analysis of data gathers akin objects together in a cluster, as opposed to objects located in disparate clusters which vary greatly from one another. The similarity rate in data is quantitatively represented by way of distance functions. Clustering methods fall into standard, fuzzy, and model-based approaches. Standard clustering methods are split into hierarchical and non-hierarchical methods. These are both referred to as hard clustering because each unit may or may not be allocated to a cluster. Fuzzy and model-based grouping methods are frequently considered to be soft because they make it simpler to assign units to clusters. In text data mining, clustering is a classification method that groups words according to the similarity of their distribution in texts, also groups documents by author or according to the similarity of their topics, etc. Comparing the usage of highfrequency features in texts is the most effective method to distinguish between the text of different authors. Clustering techniques in text document databases aim at three main concerns, namely, data sets with high dimensionality, vast databases, and a lack of clear and concise cluster descriptions, [1]. Text clustering finds various applications, [2], such as web search results clustering, automatic document organization, and social news clustering, [3], [4]. It can also be used as an intermediate step for applications such as multi-document summarization, [5], [6], real-time text summarization, [7], sentiment analysis, topic extraction, and labeling of documents. In [1], [8], based on frequent feature groups are proposed novel solutions to the problem of text clustering, with the first focusing on efficiency and accuracy, and the second on hierarchical clustering and overcoming a specific shortcoming of traditional methods. Both papers provide experimental evidence of the effectiveness of their proposed algorithms and offer insights that can be useful for future research in text clustering. Various techniques applied for author identification in different languages are presented in [9], where is concluded that no one approach is used exclusively identification: rather, researchers for author apply a variety of techniques depending on the characteristics of the understudied language, the training data set, and the feature set. Albanian language is classified as a unique branch on the IndoEuropean language family due to its unique phonological, grammatical, and lexical features and the complex syntactic structure. In the last 10 years, we have found several studies in mathematics methods and computer science for the classification of Albanian texts, for authorship attribution or author identification. Our research work is focused on developing and adapting mathematical models and statistics methods for the identification of the authors of Albanian texts with the aim of authorship attribution and the detection of plagiarism in Albanian texts, [10]. Previously we estimated the probability of finding the correct author in Albanian text classification using logistic and multinomial logistic regression models, [11], [12]. Nowadays clustering methods are applied in Albanian texts for word classification, [13], and in datasets with short comments from social networks, [14]. Several studies on Albanian text classification and categorization are made in classification texts by topics and notations of texts as positive and negative in short texts, [15], [16], [17], [18], [19]. In this paper, we present the agglomerative hierarchical clustering to classify Albanian documents by authors according to the similarity of their word frequency. We apply the agglomerative hierarchical clustering methods in a database created from 100 Albanian documents from 10 different authors. The similarity of texts is realized using cosine and Euclidian distances. The application was developed using different text mining packages in R, [20]. Considering the importance of stop words in text classification models, [21], [22], we realized the application in two cases: one with the pre-processing of the corpus by removing Albanian stop words and the other with Albanian stop words included. To increase the accuracy of classification, in this paper, we upgrade the set of Albanian stop words in R for the application of the hierarchical method as text classification. We evaluate the clustering of Albanian text by utilizing Dunn's index, thus determining the optimal clustering.

## 2 Materials and Methods

Hierarchical clustering is a powerful technique for identifying natural groupings within datasets, which can be especially useful for unsupervised text classification. Hierarchical clustering successively merges each text or document on a corpus into the default cluster based on their similarity. Similarity can be evaluated cosine similarity, Euclidean bv Manhattan distance. maximum distance. distance, etc. Hierarchical methods have the advantage of the simple interpretation of the clustering results and do

not require a prior setting of the number of clusters. The goal of clustering is to minimize the distance between the documents in the same cluster and to maximize the distance between documents in different clusters. There are two types of hierarchical methods called agglomerative and divisive methods. These techniques construct their hierarchy in the opposite direction.

Agglomerative methods start when all objects are apart then in each step two clusters are merged until only one is left. On the other hand, divisive methods start when all objects are together and in each following step, a cluster is split up, until there are all of them.

Agglomerative hierarchical clustering has been widely used in document classification, where large volumes of textual data are analyzed and categorized into groups based on their similarity. The agglomerative hierarchical clustering algorithm starts by treating each document as a separate cluster, and then iteratively merges the most similar clusters until all documents are grouped into a single cluster. A linkage criterion, such as the average linkage, complete linkage, or Ward's method, is used to determine the similarity between two clusters based on the similarities between their members. Ward's method is recognized as a highly effective technique for text clustering. This method is an agglomerative clustering technique that recursively splits the dataset into smaller subsets until each subset contains only one document. The algorithm iteratively merges the subsets that minimize the total sum of squares between each point and its corresponding centroid. This method is sensitive to outliers, as it aims to minimize the distance between data points and their respective centroids. Another successful clustering agglomerative method for text classification is the average linkage method, also referred to as the UPGMA method (unweighted pairgroup method using the average approach), which calculates the distance between two clusters as the mean of the distances between each pair of documents consisting of one member from each group. The complete linkage method tends to find uniform clusters in which the similarity between two clusters is the maximum distance between documents.

In this approach, documents are initially represented as vectors in a high-dimensional feature space, where each feature corresponds to a specific term in the document. The similarity between two documents is then measured using a distance metric. The appropriate distance for text classification is cosine distance. This method considers the angle between the document vectors and is less sensitive to outliers, as it focuses on similarity measures rather than absolute distances. Overall, the Cosine distance in average linkage method is recommended for text classification tasks as it is less sensitive to outliers and gives more weight to document similarity measures rather than distance metrics. However, these methods can be effective depending on the specific context and dataset. In this paper, we apply these methods in different distance functions to classify Albanian texts by the author.

## **3 Experimental Results and Discussion**

The data obtained from texts is regarded as a collection of terms, where a term is any string of characters separated by delimiters and may comprise one or more words. Stemming algorithms are usually employed to reduce terms to their fundamental form. Consequently, a text is transformed into a term matrix. Various text mining packages have been developed in R, [20], which include cluster analysis methods. Some of the R packages are tm, cluster, text2, word2vec, snowball, clvalid, dendextend, factoextra etc. In R we can organize the corpus in matrixes of observations and attributes. These are called document term matrices (DTM) the or transposition, term document matrices (TDM). In DTM, each row represents a document or individual corpus. The DTM columns are made of words or word groups. In the transposition matrix TDM, the word or word groups are the rows while the documents are the columns, [20]. Clustering methods applications in Albanian language texts are in datasets with short comments from social networks. Applying different methods on a dataset with three comments in the Albanian language from social networks, in [14], the authors show that the suitable algorithm is agglomerative most clustering. Using Ward's method of hierarchical clustering with Euclidian distance in [13], are defined 5 clusters of Albanian words according to the difference in frequency. To get the best clustering in [13], is created a list of 32 most the corpus frequent stop words in for preprocessing texts. In this paper, we apply Ward's, Average, Complete hierarchical and clustering methods respectively with Euclidean and Cosine distance for the classification of Albanian text by authors. We consider a corpus with 100 Albanian texts from 10 different Albanian authors. Texts in the corpus are journal papers on different topics. Each text has an average number of 1280 words. The labels of authors and texts in R, are presented in Table 1.

Text clustering techniques require multiple preprocessing steps. Initially, all non-textual elements such as symbols and punctuation are eliminated from the documents, and capital letters are converted to lower letters. Every author has a unique style of writing that stems from an unconscious habit. This is reflected in their distinctive usage of grammar, words, and punctuation which are different features for different languages. The author's style of writing is an important feature for text classification by the author in authorship attribution problems, [21].

Author	Text label	Mean
		of words number
1	1-10	1186
2	11-20	1932
3	21-30	676
4	31-40	1277
5	41-50	1201
6	51-60	1769
7	61-70	1194
8	71-80	1155
9	81-90	1048
10	91-100	1365

Table 1. Labels of texts by the author

The most frequently used words in written texts, called function words, hold a significant role as indicators of an author's style as they are employed unconsciously and can reveal significant stylistic patterns. Among the most frequent words in different languages, conjunctions, pronouns, and stop words are extensively documented as functional but noninformative words that perform a crucial role in sentence structure. Removing stop words can lead to improved accuracy in text classification models because they do not add any meaningful value in determining the category of a text, [21]. When performing text clustering based on word meaning or topics, it is essential to exclude such words to ensure accurate results. But is important to evidence that the impact of stop word removal varies based on the task, [22]. While removing stop words can lead to improved accuracy in some cases, it may not always be beneficial. For example, in sentiment analysis, removing stop words may not lead to significant improvements as stop words can be indicative of sentiment. As we mentioned above the Albanian language is a unique branch in the family of Indo-European languages, so we should consider the set of Albanian stop words and study their impact in text clustering. In this paper, we consider the corpus in two study cases, first removing the Albanian stop words list and the second not removing them. We apply and build different models to achieve the best accuracy of author classification in Albanian text. Initially, we preprocess the corpus by removing punctuation and numbers, stripping white space, and converting all capitalized letters to lowercase.

3.1 **Results in the first study** case For text classification is necessary to remove the most frequent stop words because they do not add much information to the analysis. There is no package of R where to find a list of stop words for the Albanian language. So, in [13], we created a set with the 32 most frequent stop words in the Albanian language consisting of articles, prepositions, conjunctions, pronouns, and some auxiliary verbs. In our research work, we applied the clustering methods using the list of 32 stop words in Albanian, but we achieved a low percentage of well-classified texts with the best value of 71%. To increase the classification rate, we should upgrade the Albanian stop words list. In this paper, we upgrade the set of Albanian stop words in R for the application of the hierarchical method as text classification.

We consider a set of 60 most frequent stop words among the most frequent words in the corpus presented below:

"per", "deri", "këtij", "nëse", "këto", "siç", "çdo", "ose", " disa", "është", "ishte", "kishte", "sikur", "kishin", "janë", "kanë", "ishin", "gjë", "duke", "prej", "mund", "kështu", "nga", "nuk", "kur", "kjo", "që", "dhe", "të", "në", "se", "s ë", "më", "edhe", "për", "unë", "ti", "ai", "ajo", "ne", "ju", "ata", "ato", "si", "por", "apo", "një", "ti", "t'u", "pra", "tij ", "saj", "atë", "sepse", "këtë", "tyre", "etj", "cili", "cila", "cilët"

After removing the most frequent Albanian stop words we create a term document matrix from the corpus with a matrix consisting of 19146 terms and 97% sparsity.

The complete linkage method at first was applied using cosine distance function. The optimal clustering is achieved for 16 clusters with the highest Dunn's index of 0.7521. As result, 82% of texts are well-classified by the author.



Figure 1. Comparison of cluster dendrograms of Albanian text classification with Complete linkage method in Cosine and Euclidean distance functions.



Figure 2. Comparison of cluster dendrograms of Albanian text classification with Average linkage method in Cosine and Euclidean distance functions.



Figure 3. Comparison of cluster dendrograms of Albanian text classification with Ward's method in Cosine and Euclidean distance functions.

When the Complete linkage method was applied using the Euclidian distance function the optimal clustering was achieved for 17 clusters with a classification accuracy of 82%. The method gives the same accuracy for the different distance functions. The comparison of cluster dendrograms of Albanian text classification with the complete linkage method in cosine and Euclidian distance functions is presented in Figure 1. As it is shown in Figure 1, we don't have entanglement of the branches.

Using the average linkage method in cosine distance function we get the dendrogram presented on the left of Figure 2. The best classification is in 10 clusters with an accuracy of 81%. From the dendrogram on the right of Figure 2, we notice that the average linkage method in the Euclidean distance function gives the same accuracy of 81% with the highest Dunn's index value of 0.849. As it is shown in Figure 2 the entanglement coefficient keeps the same value as in the complete linkage method but the number of clusters is reduced to 10.

Using Ward's method, we get different results for different distance functions, as it is shown in Figure 3. Applying the cosine distance function, we get the dendrogram on the left of Figure 3. The best classification with 16 clusters is achieved with the highest Dunn index value of 0.72 with the overall percentage of text classification by the author from the clusters with a value of 87%. Ward's method in Euclidean distance function gives the dendrogram on the right of Figure 3. The best classification with 16 clusters is achieved with the highest Dunn index value of 0.849 and the overall percentage of text classification by the author from the clusters is with a value of 84%. Although Ward's method, in both distance functions, determines the same number of clusters, the entanglement coefficient value of 0.57 indicates a high crossing alignment of texts. This explains the difference in the accuracy of text classification.

Results achieved in different agglomerative hierarchical clustering on the classification of Albanian texts by authors according to the similarity of their word frequency are summarized in Table 2.

As conclusion, the optimal clustering for Albanian texts in the case of removing the most frequent stop words is achieved by Ward's method in cosine distance with the highest Dunn index value of 0.72 and with the highest accuracy 87%.

Table 2.	The results from the application in R
	for the first study case.

Method	Distance	Percentage of texts well classified by author	Number of clusters	Highest Dunn's index
Complete	Euclidean	82	17	0.8741
linkage	Cosine	82	16	0.7527
Average	Euclidean	81	10	0.8893
linkage	Cosine	81	10	0.7909
Ward	Euclidean	84	16	0.8490
	Cosine	87	16	0.7209

#### **3.2** Results in the second study case

In this case, we apply agglomerative methods in the corpus without removing the stop words list we mentioned above. After the preprocessing of the texts, we created a term document matrix from the corpus with a matrix consisting of 19283 terms and 97% sparsity.

Initially, we apply the complete linkage using the cosine distance function, as a result, 62% of texts are well-classified by the author. We get the same result of text classification when we apply the complete linkage using the Euclidian distance function. Using the complete linkage method, we compare the dendrograms in cosine and Euclidean distance functions respectively in Figure 4. As it is shown in Figure 4, from the dendrograms, we notice that the entanglement value is 0. We can't determine the optimal clustering because Dunn's index increases when the number of clusters is increased. As result, 62% of texts are well-classified by the author.

Figure 5, presents the comparison of dendrograms for the average linkage method. Using the average linkage method in the cosine distance function we get the dendrogram presented on the left of Figure 5. The best classification is in 10 clusters with an accuracy of 64% with the highest Dunn index value of 0.4642. As it is shown from the dendrogram on the right of Figure 5, the average linkage method in Euclidean distance function gives the same accuracy of 66% with the highest Dunn index value of 0.7005. The negligible entanglement coefficient indicates the best result when we use the Euclidian distance function.



Figure 4. Comparison of cluster dendrograms of Albanian text classification with Complete linkage method in Cosine and Euclidean distance functions.



Figure 5. Comparison of cluster dendrograms of Albanian text classification with Average linkage method in Cosine and Euclidean distance functions.



Figure 6. Comparison of cluster dendrograms of Albanian text classification with Ward's method in Cosine and Euclidean distance functions.

Using Ward's method, we get different results for different distance functions as shown in Figure 6. Applying the cosine distance function, we get the dendrogram on the left of Figure 6. The best classification with 15 clusters is achieved with the highest Dunn index value of 0.72 with the overall percentage of text classification by the author from the clusters with a value of 75%. Ward's method in Euclidean distance function gives the dendrogram on the right of Figure 6. The best classification with 16 clusters is achieved with the highest Dunn index value of 0.849. As is seen in Figure 6, the overall percentage of text classification by the author from the clusters is with a value of 75%. Although Ward's method, in both distance functions, determines the same accuracy of text classification, the small value of the entanglement coefficient of 0.0083 indicates the difference in cluster numbers.

Results achieved in different agglomerative hierarchical clustering on the classification of Albanian texts by authors according to the similarity of their word frequency are summarized in Table 3. As conclusion, the optimal clustering for Albanian texts in the second study case is achieved by Ward's method in cosine distance with the highest accuracy of 75%.

In the second study case						
Method	Distance	Percentage	Number	Highest		
		of texts well	of	Dunn's		
		classified by	clusters	index		
		author				
Complete	Euclidean	62	-	0.38-		
linkage				1.06		
	Cosine	62	-	0.62-		
				1.03		
Average	Euclidean	66	14	0.7005		
linkage	Cosine	64	10	0.4642		
Ward	Euclidean	75	16	0.8490		
	Cosine	75	15	0.7209		

 Table 3. Results of agglomerative hierarchical clustering in the second study case

#### 3.3 Discussion

Results presented in Table 2 and Table 3 show that we get the accuracy of classification at least 80% in the first case of study. From these results, we conclude that removing the most frequent stop words improves the accuracy of Albanian text classification. In the first case, complete and average linkage methods give the same results for the two distance functions. In the second one, it's not possible to determine the optimal clustering for the complete linkage method because Dunn's index can only keep increasing until all the texts are separated into 100 clusters. In [12], using the logistic regression model as a classification method to estimate the probability of finding the correct author in Albanian text was concluded that the multinomial logistic regression model for Albanian text has more advantages than the logistic regression model with the highest overall correct predicted probability 0.738. By removing the stop words from the corpus, the cosine distance function proved to be more efficient, while in the second case, by including the stop words in the corpus, the Euclidean distance function was more efficient. Comparing the achieved outcomes, the best results for text classification by the author are Albanian obtained using Ward's method in the cosine distance function, increasing the accuracy to 87%.

# 4 Conclusion

Using agglomerative hierarchical clustering methods in a corpus with 100 Albanian texts of 10 different authors and the rich packages of text mining in R we realized different classifications of texts by the author according to the similarity of their word frequency. In the tm package of R, we have successfully implemented the upgrade list with the 60 most frequented stop words in the Albanian language. The optimal clustering for each method was determined using Dunn's agglomerative hierarchical index. Applying clustering methods in Albanian texts in two cases we conclude that removing the most frequent stop words improves the accuracy in Albanian text classification. The best classification is achieved using Ward's method of cosine distance. The best model evaluated with the maximum value of Dunn's index of 0.7209, separated the Albanian texts into 16 different clusters based on the frequency of words. From the clusters, we estimate the overall percentage of text classification by author, with a value of 87%. Our research showcases the capability of agglomerative hierarchical clustering techniques in identifying the authors of Albanian texts. Although we achieved a promising accuracy rate of 87%, there

are opportunities for improvement and expansion through larger datasets. Although we improved the outcomes by incorporating a list of stop words for the Albanian language, the selection was limited to our small data set. The future work will focus on the selection of stop words which may still be subject to further refinement to achieve better results. Extensive research needs to be done to improve the models used and to explore statistical methods of machine learning approaches.

References:

- [1] Beil, F., Ester, M., Xu, X.: Frequent term-based text clustering. In: KDD. (2002)
- [2] Aggarwal CC, Zhai C(2012) A survey of text clustering algorithms. Mining text data. Springer, New York, pp 77–128.
- [3] Xia Y, Tang N, Hussain A, Cambria E (2015) Discriminative biterm topic model for headlinebased social news clustering. In: The twentyeighth international flairs conference, pp 311– 316
- [4] Yan X, Guo J, Lan Y, Cheng X (2013) A biterm topic model for short texts. In: Proceedings of the 22nd international conference on World Wide Web. ACM, pp 1445–1456
- [5] Saggion H, Poibeau T (2013) Automatic text summarization: past, present and future. Multisource, multilingual information extraction and summarization. Springer, New York, pp 3–21
- [6] Turpin A, Tsegay Y, Hawking D, Williams HE (2007) Fast generation of result snippets in web search. In: Proceedings of the 30th annual international ACM SIGIR conference on research and development in information retrieval, pp 127– 134
- [7] Liu CY, Chen MS, Tseng CY (2015) Incrests: towards real-time incremental short text summarization on comment streams from social network services.IEEE Trans Knowl Data Eng 27(11):2986–3000
- [8] Fung, B. C. M., Wang, K., Ester, M.: Hierarchical Document Clustering Using Frequent Itemsets. In SDM. (2003)
- [9] Digamberrao, K. S., & Prasad, R. S. (2018). Author Identification on Literature in Different Languages: A Systematic Survey. 2018 International Conference on Advances in Communication and Computing Technology, ICACCT 2018, 174–181. https://doi.org/10.1109/ICACCT.2018.8529635

- [10] H. Paci, E. Kajo, E. Trandafili, I. Tafa and D. Salillari, "Author Identification in Albanian Language," 2011 14th International Conference on Network-Based Information Systems, Tirana, Albania, 2011, pp. 425-430, doi: 10.1109/NBiS.2011.71.
- [11] Salillari, D., & Prifti, L. (2016). A multinomial logistic regression model for text in Albanian language. JOURNAL OF ADVANCES IN MATHEMATICS, 12(7), 6407–6411. <u>https://doi.org/10.24297/jam.v12i7.5486</u>
- [12] Salillari, D., & Prifti, L. (2016). Comparison Study of Logistic Regression Model for Albanian Texts. JOURNAL OF ADVANCES IN MATHEMATICS, 12(9), 6572–6575. <u>https://doi.org/10.24297/jam.v12i9.127</u>
- [13] Denisa Salillari, Luela Prifti "Cluster analysis and its application in Albanian texts" Proceedings book of International Conference on ICEAS2022 17-18 November 2022.
- [14] Mërgim H. HOTI, Jaumin AJDARI, "Unsupervised Clustering of Comments Written in Albanian Language" International Journal of Advanced Computer Science and Applications (IJACSA), 12(8), 2021.
- [15] A. Kadriu and L. Abazi, "A comparison of algorithms for text classification of Albanian news articles," Entrenova-Enterprise Research Innovation Conference, vol. 3, no. 1, pp. 62–68, 2017.
- [16] E. Trandafili, N. Kote, and M. Biba, "Performance evaluation of text categorization algorithms using an Albanian corpus," in Advances in Internet, Data & Web Technologies, 2018, pp. 537–547.
- [17] Biba, M. and Mane, M. (2014) Sentiment Analysis through Machine Learning: An Experimental Evaluation for Albanian. In: Thampi, S., Abraham, A., Pal, S. and Rodriguez, J., Eds., Recent Advances in Intelligent Informatics, Springer International Publishing, Cham, 195-203. https://doi.org/10.1007/978-3-319-01778-5\_20
- [18] Skenduli, M.P., Biba, M. (2020). Classification and Clustering of Emotive Microblogs in Albanian: Two User-Oriented Tasks. In: Appice, A., Ceci, M., Loglisci, C., Manco, G., Masciari, E., Ras, Z. (eds) Complex Pattern Mining. Studies in Computational Intelligence, vol 880. Springer, Cham. https://doi.org/10.1007/978-3-030-36617-9\_10
- [19] A. Kadriu, L. Abazi, and H. Abazi, "Albanian text classification: Bag of words model and word analogies," Business Systems Research Journal, vol. 10, no. 1, pp. 74–87, Apr. 2019, doi: 10.2478/bsrj-2019-0006.0
- [20] Ted Kwartler, Text Mining in Practice with R, John Wiley & Sons Ltd, 2017.
- [21] Stamatatos, E., Tschnuggnall, M., Verhoeven, B., Daelemans, W., Specht, G., Stein, B., & Potthast, M. (2016). Clustering by authorship

within and across documents. In Working Notes Papers of the CLEF 2016 Evaluation Labs. CEUR Workshop Proceedings/Balog, Krisztian [edit.]; et al. (pp. 691-715).

[22] Panicheva, P., Litvinova, O., Litvinova, T. (2019). Author Clustering with and Without Topical Features. In: Salah, A., Karpov, A., Potapova, R. (eds) Speech and Computer. SPECOM 2019. Lecture Notes in Computer Science(), vol 11658. Springer, Cham. https://doi.org/10.1007/978-3-030-26061-3\_36

#### Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

#### Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

#### **Conflict of Interest**

The authors have no conflicts of interest to declare that are relevant to the content of this article.

# Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0 https://creativecommons.org/licenses/by/4.0/deed.en

\_US