I. J. Mathematical Sciences and Computing, **2022, 3, 21-29**
Published Online August 2022 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijmsc.2022.03.02

Outlier Detection Algorithm Based on Fuzzy C-Means and Self-organizing Maps Clustering Methods

Mesut Polatgil
Şarkışla School of Applied Sciences /Computer Science, Sivas, 58050, Turkey
E-mail: mesutpolatgil@cumhuriyet.edu.tr

Received: 13 March 2022; Accepted: 25 May 2022; Published: 08 August 2022

Abstract: Data mining and machine learning methods are important areas where studies have increased in recent years. Data is critical for these areas focus on inferring meaningful conclusions from the data collected. The preparation of the data is very important for the studies to be carried out and the algorithms to be applied. One of the most critical steps in data preparation is outlier detection. Because these observations, which have different characteristics from the observations in the data, affect the results of the algorithms to be applied and may cause erroneous results. New methods have been developed for outlier detection and machine learning and data mining algorithms have been provided with successful results with these methods. Algorithms such as Fuzzy C Means (FCM) and Self Organization Maps (SOM) have given successful results for outlier detection in this area. However, there is no outlier detection method in which these two powerful clustering methods are used together. This study proposes a new outlier detection algorithm using these two powerful clustering methods. In this study, a new outlier detection algorithm (FUSOMOUT) was developed by using SOM and FCM clustering methods together. With this algorithm, it is aimed to increase the success of both clustering and classification algorithms. The proposed algorithm was applied to four different datasets with different characteristics (Wisconsin breast cancer dataset (WDBC), Wine, Diabetes and Kddcup99) and it was shown to significantly increase the classification accuracy with the Silhouette, Calinski-Harabasz and Davies-Bouldin indexes as clustering success indexes.

Index Terms: Outlier detection, Fuzzy C Means, Self-Organization Maps, Silhouette, Calinski-Harabasz, Davies-Bouldin.

1. Introduction

With the increasing population and developing technology, there are serious increases in the amount of data obtained. Algorithms such as data mining and machine learning focus on inferring meaningful conclusions from this data. In order to infer meaningful conclusions from the data, the preparation of the data, that is, the pre-processing, is a very important step. One of the most critical points of this stage is outlier detection. Outlier observations exhibit different characteristics when compared to other observations in the data set. If it is viewed through a two-dimensional or three-dimensional graphic, it is seen that these observation points are at a very different point from other data in terms of location [1]. Therefore, in different areas outlier observations can also be expressed with words such as anomalies, discordant, noise, error or surprise. The presence of these outlier observations in the data set may be caused by factors such as the device used to measure, human, environment, sampling and experimental errors [2,3].

Although outlier observations are thought to be caused by an error or inaccuracy, in some cases they may also be observations containing meaningful information [4]. Detecting these observations can be critical for some disciplines. For example, outlier detection has an important role in detecting data that is different from the data of healthy individuals in medical applications and taking action immediately [5] or detecting actions such as credit card fraud [6].

There are three different approaches in which a data is outlier. These are Global, Collective and Contextual approaches. In the global approach, data points that are quite different from other observations in the data set can be outliers, while a group of data can be outliers in a collective outlier. Contextual outlier observations refer to data that are likely to be outlier observations in terms of time, space and place. Outlier detection methods are classified in four different ways in the literature: distance-based, cluster-based, density-based and statistical-based [4].

Among these ways, clustering-based have been widely used in outlier detection in recent years and new algorithms have been developed. In the study [2], o utlier detection was performed by using K means, K++ means and FCM algorithms together. In the study [7], outlier detection was done with data mining. Studies using only Fuzzy clustering are also available in the literature [8,9,10,11].

Outlier detection was performed using the SOM algorithm in the literature [12,13]. The SOM method and the PSO method were used together in the detection of credit card fraud [14], Robust self-organization with M-estimators [15], Creation of Text Document Matrices and Visualization by Self-Organizing Map [16], Spatial outlier detection based on iterative self-organizing learning model [17], Multivariate outlier detection based on self-organizing map and adaptive nonlinear map [18], SOM neural network compares with Fuzzy c-means, K-means and traditional hierarchical clustering algorithms [19], how to use Self-organizing maps for outlier detection [20], how to visualization of Outliers via Self-Organizing Maps [21,22].

However, there is no outlier detection method in which these two powerful clustering methods are used together. This study proposes a new outlier detection algorithm using these two powerful clustering methods. In the study [3], in which many clustering methods, especially both FCM and SOM algorithms, were evaluated together, outlier detection was made with the voting system.

In this study, it is aimed to improve machine learning results by using two powerful methods such as FCM and SOM together in the outlier detection task. Although these two algorithms were used in the Outlier detection task, they were not used together. The limitation of the proposed method is that it is necessary to know how to use both algorithms for outlier detection.

In this study, a new algorithm named FUSOMOUT is proposed, which combines FCM and SOM clustering methods to detect outliers. By applying the FCM algorithm to the data set, the observations were assigned to the clusters with certain probabilities. The median value was found by arranging the highest probability values for each observation, and a threshold value was determined by going back 25% from this median value. Observations with probability values below this threshold value were determined as outlier data and stored. In order to clarify the outlier status of these data, the SOM clustering method was applied to the data set and the average observation value for each node was taken as a basis by using the Hit matrix values in the SOM map. Nodes with observations below this value were determined as outliers. Then, according to the results of both FCM and SOM algorithms, the stored observation points with the possibility of being outliers were determined as outliers. It was run on four different datasets to show how this outlier detection algorithm affects clustering and classification results. These data sets; WDBC, Wine, Diabets and Kddcup99 are determined as datasets. In the selection of these data sets, the fact that the data sets have different attributes and observation numbers and the use of studies that propose new methods for outlier detection were taken into consideration. The remainder of the study is organized as follows, respectively. In the chapter, FCM and SOM clustering methods and working methods are explained. In the next section, the proposed algorithm and working steps are shown, and in the next section, it is run on four different data sets and the results are presented.

1.1. Fuzzy C Means

Clustering methods are generally evaluated in two categories as hard and soft clustering. Hard clustering methods assume that each observation belongs to only one cluster. Fuzzy C Means algorithm is a soft clustering method. In this method, instead of belonging to a single cluster, observations are assigned to more than one cluster with a membership degree [23,24]. In the algorithm, the objective function given in (1) is tried to be minimized.

$$J = \sum_{i=1}^{N}\sum_{k=1}^{C}\mu_{ik}^{m}\left\|x_i - c_k\right\| \tag{1}$$

μ_{ik} indicates the degree of belonging of the observations to the clusters, m the degree of fuzzy, xi the observation points and Ck the cluster centers.

The algorithm works as follows:

- The number of clusters is selected
- Cluster membership values are determined randomly.
- The algorithm is repeated if the difference in membership values between the two iterations is not less than €.
- Cluster centers are calculated.
- The membership of the observations to the cluster centers is calculated.

Cluster centers are calculated as given in (2)

$$c_k = \frac{\sum_{i=1}^{N}\mu_{ik}^{m} * x}{\sum_{i=1}^{N}\mu_{ik}^{m}} \tag{2}$$

The membership values of the observations belonging to the clusters are calculated as given in (3):

$$\frac{1}{\mu_{ij}} = \sum_{k=1}^{c} \left(\frac{x_i - C_j}{x_i - C_k} \right)^{\frac{2}{m-1}} \tag{3}$$

1.2. SOM

SOM is a special case of artificial neural networks developed for the clustering method [22]. As seen in Figure 1, while the input layer contains the features in the data set, the output layer is usually in the form of a grid of two-dimensional nodes.

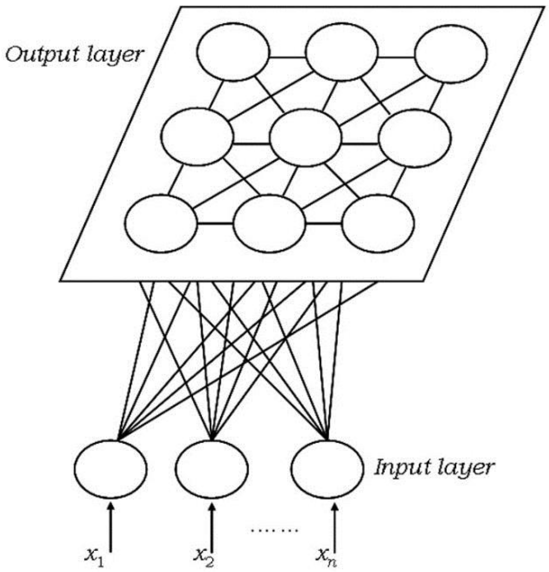


Fig. 1. SOM structure

Each observation in the data set is placed at the nearest node and this node is called the winning node. The winning node is calculated as given in (4):

$$win(t) = \arg\min\{\| x - w_k \|\}(t), k \in \{1, 2, 3, ..., m\} \tag{4}$$

Where x input data, w_k output layer k. Node, m is the number of nodes in the output layer and $\|x - w_k\|$ Euclidean distance.

The working steps of the algorithm are as follows:

- Weights are randomly assigned to the nodes in the output layer.
- For each observation in the dataset, the closest node, that is, the winning node, is found in the output layer.
- The weights of the winning node and neighboring nodes are updated.
- As long as the number of iterations is less than the number, the algorithm continues from the 2nd step.

Updating the weights of the winning node and neighboring nodes is done according to (5):

$$\text{e } w_{ij}(t+1) = w_{ij}(t) + h_{ij}^{w}(x_n - w_{ij}(t)) \tag{5}$$

Where H_{ij} neighborhood function and is used to determine the nodes whose weights will be updated together with the winning node in the output layer [13].

2. Method

2.1. Proposed Algorithm

FUSOMOUT, a new outlier detection algorithm based on FCM and SOM clustering methods, is proposed from this study. Outlier detection was performed by using FCM and SOM algorithms together. Then, these data were deleted and clustering and classification operations were performed on the datasets. After removing the outlier data, clustering and classification methods were applied again, and the results were shown.

The working steps of the algorithm are shown in Figure 2.

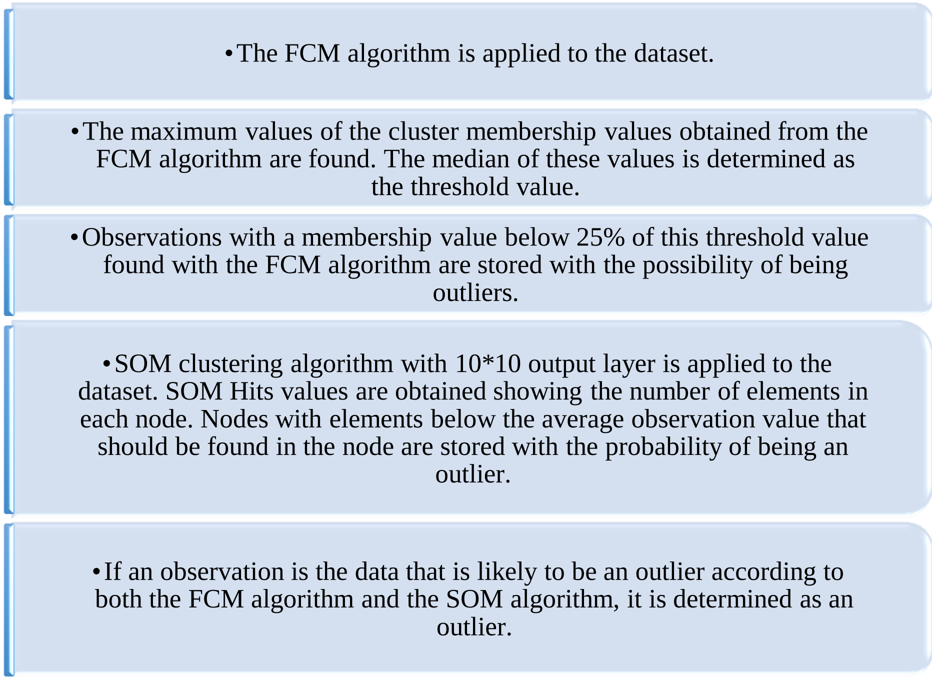


Fig. 2. *Working steps of the proposed FUSOMOUT algorithm.*

2.2. Dataset

In order to test the performance of the proposed algorithm, a selection was made from the datasets in the studies in which new algorithms for outlier detection were included in the literature in recent years. In this context, four different verses were selected based on studies using both FCM and SOM clustering methods [1,2,8]. Descriptive information about these selected datasets is given in Table 1.

Table 1. Datasets used in the study

Dataset	No. of samples	No. Of attributes	No. Of class
WDBC	699	9	2
Wine	178	13	3
Diabets	768	8	2
Kddcup99	494.021	41	23

2.3. Clustering Index

The clustering results obtained by applying the FUSOM algorithm proposed in the study to these data sets were calculated with three different indexes. These are the Silhouette, Calinski-Harabasz and Davies-Bouldin indexes.

2.4. Silhouette Index

This clustering validation criterion shows how close the observations in the data set are to the elements in its own cluster compared to other clusters [25]. Silhouette index calculation method is given in (6):

$$S(i) = \frac{\left(b(i) - a(i)\right)}{\max\{a(i), b(i)\}} \tag{6}$$

Here b(i), the average distance of the i. observation point to the elements in the cluster closest to the cluster, a(i), It shows the average distance of the i. observation point to the elements in its cluster. The Silhouette value is calculated separately for each observation value, and the Silhouette value of the entire data set is the average of the Silhouette values of each observation data. A high index indicates good clustering results. The Silhouette index provides a graphical representation of the calculated values, making it widely used compared to other clustering indexes.

2.5. Calinski-Harabasz Index

The Calinski-Harabasz index is a clustering measure based on the rate of variance. It is calculated according to the ratio of the inter-cluster variance to the intra-cluster variance as shown in (7) [26]:

$$CH = \frac{SS_B}{SS_w} * \frac{(N-K)}{(K-1)} \tag{7}$$

SS_b is the variance inter-clusters, SS_w is the sum of intra-cluster variance values, K is the number of clusters and N is the number of observations in the dataset. Since the high inter-cluster variance value and low intra-cluster variance value will increase clustering success, it is desirable to have a high Calinski-Harabasz index value.

2.6. Davies-Bouldin Index

High Silhouette and Calinski-Harabasz index values indicate a good result, while a low Davies-Bouldin index value indicates a good clustering result. This index is calculated as given in (8) [25]:

$$DB = \frac{1}{C}\sum_{i=1}^{c} d_i, \textit{where } d_i = \max\{\frac{S_i + S_j}{M_{ij}}\} \tag{8}$$

Where S is the measure of variance within the cluster, C number of cluster ve M_{ij} is the distance between clusters i and j.

3. Results

In the study, first of all, clustering analysis was performed without any operation on the dataset. Then, the FUSOMOUT algorithm was applied on the data set and outlier detection was made. After the data detected as outliers were deleted, clustering was performed again on the data set. The results were given in Table 2.

Table 2. Clustering index values performed with the proposed algorithm

Dataset/ İndex	Silhouette		Calinski-Harabasz		Davies-Bouldin	
	Before	After	Before	After	Before	After
WDBC	0.594	0.754	1031	1740	0.76	0.52
Wine	0.655	0.71	505	610	0.48	0.41
Diabets	0.56	0.6	3024	4025	0.57	0.52
Kddcup99	0.598	0.723	541	863	1.34	1.16

In order to graphically show the success of the proposed FUSOMOUT algorithm on clustering indexes, its success on datasets only in terms of Silhouette Index value is shown in Figure 3. However, according to the findings of all three methods, it is seen that the proposed method increases the clustering success results.

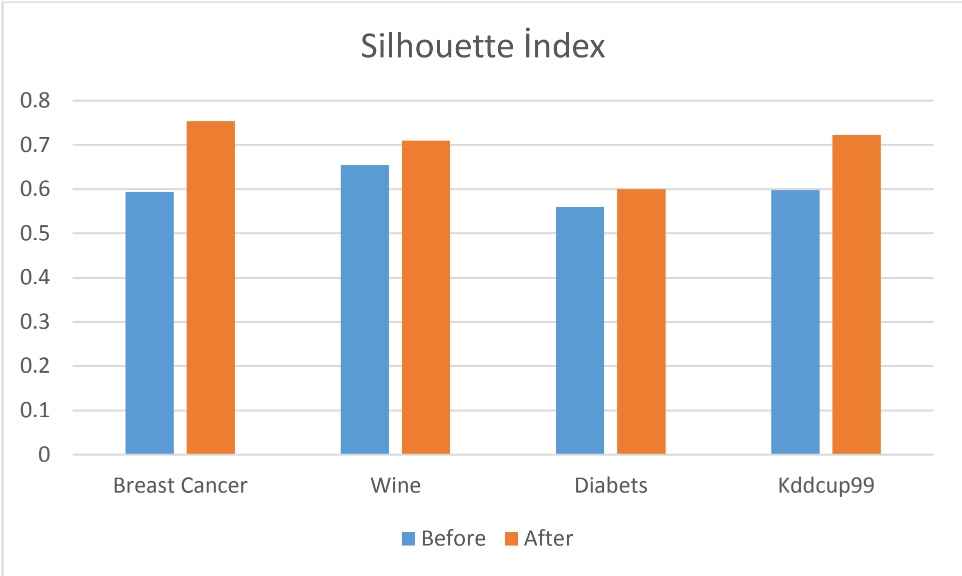
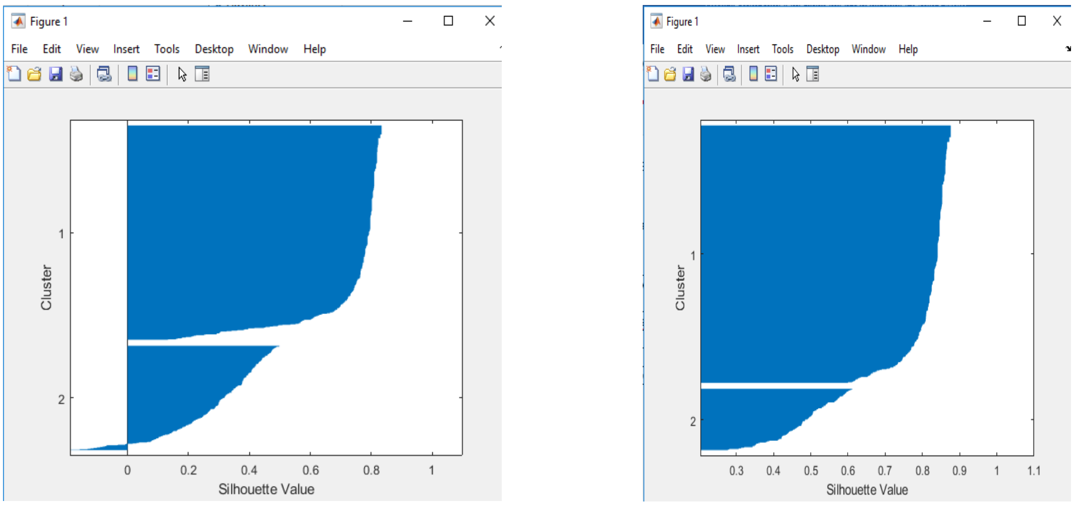


Fig. 3. Effect of FUSOMOUT algorithm on Silhouette value

It is possible to calculate the silhouette index value for each observation value and to display it graphically. Silhouette index graph of the observations in the WDBC dataset is given in Figure 4.



(a) (b)

Fig. 4. Silhouette chart of WDBC Dataset. (a) Orjinal dataset, (b) after applied FUSOMOUT

The number of data deleted from the datasets with the FUSOMOUT algorithm and their ratios are given in Table 3. It has been observed that outlier detection is generally performed with a rate below 10% of the dataset.

Table 3. Number of observations extracted by outlier detection

	Num of Observation	Deleted	Ratio
WDBC	699	51	%7.29
Wine	178	13	%7.86
Diabets	768	53	%6.9
Kddcup99	494.021	48950	%9.9

The success of the FUSOMOUT outlier detection algorithm proposed in the study has also been tested in supervised learning approaches. In this context, classification processes were performed on the four data sets before and after applying the FUSOMOUT algorithm, and the results are given in Table 4. Since it is the primary goal to show the success of the proposed FUSOMOUT algorithm here, the tests were carried out with a 5 k-fold cross-validation test without parameter tuning in the algorithms.

Table 4. Classification accuracy values by deleting outlier observations

	KNN		LR		MLP	
	Before	After	Before	After	Before	After
WDBC	0.934	0.984	0.959	0.983	0,934	0.982
Wine	0.929	0.958	0.977	0.976 2	0.983	0.97
Diabets	0.734	0.778	0.719	0.756	0.72	0.771
Kddcup99	0.853	0.949	0.901	0.968	0.919	0.967

4. Discussion

FUSOMOUT, a new outlier detection algorithm based on FCM and SOM clustering methods, is proposed from this study. Thus, a better outlier detection method has been developed by using the success of both FCM and SOM methods together.

According to the results of Table 2, the proposed FUSOMOUT algorithm in the study has significantly increased the clustering index values. In the WDBC dataset, it increased the Silhouette Index value by approximately 27%, increased the Calinski-Harabasz index value by 69% and decreased the Davies-Bouldin Index value by approximately 32%. In the Kddcup99 dataset, it increased the Silhouette Index value by 21%, increased the Calinski-Harabasz index value by 60% and decreased the Davies-Bouldin Index value by approximately 14%. FUSOMOUT algorithm showed a significant improvement especially in Calinski-Harabasz index value in these index values which show clustering success.

[Kannan et al., 2014], showed that the KEFCM and KFCM algorithms he proposed in his study increased the clustering success on WDBC and Wine datasets. However, only the Silhouette Index value was used in the study. Similarly, [28] used Wine and WDBC datasets in their study, but gave only Silhouette value. Dik et al., (2020), applied the algorithm he developed on the WDBC dataset in his study, but stated that it was not an outlier. The results of the studies that measure the clustering success and give the clustering index values related to the WDBC and Kddcup99 data sets were compared and the results are given in Table 5 [2]. It was seen that the FUSOMOUT algorithm was more successful on clustering index values in both data sets where the Ensemble method used in this study was applied.

Table 5. Comparison of FUSOMOUT and Ensemble methods with WBBC and Kaddcup99 datasets

	Silhouette		Calinski-Harabasz		Davies-Bouldin	
	FUSOMOUT	Ensemble	FUSOMOUT	Ensemble	FUSOMOUT	Ensemble
WDBC	0.754	0.655	1740	1162	0.52	0.617
Kddcup99	0.753	0.763	863	191	1.06	1.06

In terms of the Wine data set, the K-means algorithm optimized the initial state, while the Silhouette value improved to 0.5929, while the FUSOMOUT algorithm increased it to 0.755 [29]. In the study performed on the Wine dataset, the Silhouette value was found to be 0.627 [28]. Although there is a study on clustering improvement related to the Wine dataset, there is also a study in the literature that cannot be compared because clustering index values are not given [10]. In addition, the FUSOMOUT algorithm also improved the three clustering index values on the Wine dataset.

In the study, in which the outlier detection method developed on the diabetes data set was applied, results were not given in terms of clustering index values [1]. In addition, the FUSOMOUT algorithm also improved on the Diabetes dataset in terms of three clustering index values.

In terms of the results given in Table 4, the FUSOMOUT algorithm showed significant improvements in classification accuracy on classification algorithms when outlier values were removed from the data set.

5. Conclusion

Since all inferences are made from the data set in data mining and machine learning methods, making the data suitable for analysis is a very critical process. One of the most important stages of this process is outlier detection. Although new methods based on FCM and SOM algorithms have been proposed for outlier detection, there is no algorithm that uses these two methods together. In this study, a new outlier method called FUSOMOUT is proposed. The proposed method has been tested on four different data sets and the results are given in terms of three clustering index values and classification accuracy. It has been shown that the proposed algorithm makes significant improvements in terms of clustering index values and classification accuracy. The use of the FUSOMOUT algorithm in the data pre-processing in the studies to be carried out in the field of data science will make serious improvements on the results in the studies to be carried out.

In future studies, it is considered that the FUSOMOUT algorithm will be tested on new data sets. Since the results are examined on the basis of clustering and classification in the study, it is planned to examine the regression values in future studies. In addition, since the new algorithms developed for the FCM algorithm increase on the clustering index values, the effect on the results can be examined by applying the FUSOMOUT algorithm before applying these algorithms [27].

References

[1] A. Dik, K. Jebari, A. Ettouhami, An Improved Robust Fuzzy Algorithm for Unsupervised Learning, Journal of Intelligent Systems. 29 (2020) 1028–1042. https://doi.org/10.1515/JISYS-2018-0030.
[2] A. Saha, A. Chatterjee, S. Ghosh, N. Kumar, R. Sarkar, An ensemble approach to outlier detection using some conventional clustering algorithms, Multimedia Tools and Applications. (2020). https://doi.org/10.1007/S11042-020-09628-5.
[3] M. Vaghefi, K. Mahmoodi, M. Akbari, A comparison among data mining algorithms for outlier detection using flow pattern experiments, Scientia Iranica. 25 (2018) 590–605. https://doi.org/10.24200/SCI.2017.4182.
[4] A. Smiti, A critical overview of outlier detection methods, Computer Science Review. 38 (2020) 100306. https://doi.org/10.1016/j.cosrev.2020.100306.
[5] Z. Nazari, D. Kang, Evaluation of Multivariate Outlier Detection Methods with Benchmark Medical Datasets, INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND NETWORK SECURITY. 18 (2018) 36–43. https://www.webofscience.com/wos/woscc/full-record/WOS:000432495400006 (accessed August 22, 2021).
[6] F. Carcillo, Y.A. Le Borgne, O. Caelen, Y. Kessaci, F. Oblé, G. Bontempi, Combining unsupervised and supervised learning in credit card fraud detection, Information Sciences. 557 (2021) 317–331. https://doi.org/10.1016/J.INS.2019.05.042.
[7] P. Verma, M. Sinha, S. Panda, Fuzzy c-Means Clustering-Based Novel Threshold Criteria for Outlier Detection in Electronic Nose, IEEE Sensors Journal. 21 (2021) 1975–1981. https://doi.org/10.1109/JSEN.2020.3020272.
[8] K. Vembandasamy, T. Karthikeyan, Novel outlier detection in diabetics classification using data mining techniques | Request PDF, International Journal of Applied Engineering Research. 11 (2016) 1400–1415. https://www.researchgate.net/publication/298714562_Novel_outlier_detection_in_diabetics_classification_using_data_mining_techniques (accessed August 8, 2021).
[9] H. Tanatavikorn, Y. Yamashita, Fuzzy treatment method for outlier detection in process data, Journal of Chemical Engineering of Japan. 49 (2016) 864–873. https://doi.org/10.1252/JCEJ.16WE042.
[10] A. Dik, K. Jebari, A. Bouroumi, A. Ettouhami, A new fuzzy clustering by outliers, Journal of Engineering and Applied Sciences. 9 (2014) 372–377. https://doi.org/10.3923/JEASCI.2014.372.377.
[11] M.B. Al-Zoubi, A. Al-Dahoud, A.A. Yahya, New outlier detection method based on fuzzy clustering, WSEAS Transactions on Information Science and Applications. 7 (2010) 681–690.
[12] P. Stefanovic, O. Kurasovay, Outlier detection in self-organizing maps and their quality estimation, Neural Network World. 28 (2018) 105–117. https://doi.org/10.14311/NNW.2018.28.006.
[13] T.Y. Christyawan, W.F. Mahmudy, Text Classification and Visualization on News Title Using Self Organizing Map, in: 3rd International Conference on Sustainable Information Engineering and Technology, SIET 2018 - Proceedings, Institute of Electrical and Electronics Engineers Inc., 2018: pp. 332–336. https://doi.org/10.1109/SIET.2018.8693189.
[14] S. Arora, D. Kumar, Hybridization of SOM and PSO for detecting fraud in credit card, International Journal of Information Systems in the Service Sector. 9 (2017) 17–36. https://doi.org/10.4018/IJISSS.2017070102.
[15] E. López-Rubio, E.J. Palomo, E. Domínguez, Robust self-organization with M-estimators, Neurocomputing. 151 (2015) 408–423. https://doi.org/10.1016/j.neucom.2014.09.024.
[16] P. Stefanovič, O. Kurasova, Creation of Text Document Matrices and Visualization by Self-Organizing Map, Information Technology and Control . 43 (2014) 37–46. https://doi.org/10.5755/J01.ITC.43.1.4299.
[17] Q. Cai, H. He, H. Man, Spatial outlier detection based on iterative self-organizing learning model, Neurocomputing. 117 (2013) 161–172. https://doi.org/10.1016/j.neucom.2013.02.007.
[18] X. Yan, Multivariate outlier detection based on self-organizing map and adaptive nonlinear map and its application, Chemometrics and Intelligent Laboratory Systems. 107 (2011) 251–257. https://doi.org/10.1016/j.chemolab.2011.04.007.
[19] S.A. Mingoti, J.O. Lima, Comparing SOM neural network with Fuzzy c-means, K-means and traditional hierarchical clustering algorithms, European Journal of Operational Research. 174 (2006) 1742–1759. https://doi.org/10.1016/j.ejor.2005.03.039.
[20] A. Muñoz, J. Muruzábal, Self-organizing maps for outlier detection, Neurocomputing. 18 (1998) 33–60. https://doi.org/10.1016/S0925-2312(97)00068-4.
[21] J. Muruzábal, A. Muñoz, J. Muruzabal, A. Munoz, On the Visualization of Outliers via Self-Organizing Maps, Journal of Computational and Graphical Statistics. 6 (1997) 355. https://doi.org/10.2307/1390741.
[22] T. Kohonen, Self-organized formation of topologically correct feature maps, Biological Cybernetics 1982 43:1. 43 (1982) 59–69. https://doi.org/10.1007/BF00337288.
[23] J.C. Bezdek, Models for Pattern Recognition, Pattern Recognition with Fuzzy Objective Function Algorithms. (1981) 1–13. https://doi.org/10.1007/978-1-4757-0450-1_1.
[24] J.C. Dunn, A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, Journal of Cybernetics. 3 (1973) 32–57. https://doi.org/10.1080/01969727308546046.
[25] S.P. Petrović, A Comparison Between the Silhouette Index and the Davies-Bouldin Index in Labelling IDS Clusters, (2006).
[26] Y. Kwon, K. Kang, C. Bae, Unsupervised learning for human activity recognition using smartphone sensors, Expert Systems with Applications. 41 (2014) 6067–6074. https://doi.org/10.1016/J.ESWA.2014.04.037.
[27] S.R. Kannan, S. Ramthilagam, R. Devi, T.P. Hong, Fuzzy C-means in finding subtypes of cancers in cancer database, Journal of Innovative Optical Health Sciences. 7 (2014) 1450018. https://doi.org/10.1142/S1793545814500187.

[28] A. Triayudi, I. Fitri, Comparison of parameter-free agglomerative hierarchical clustering methods, ICIC Express Letters. 12 (2018) 973–980. https://doi.org/10.24507/ICICEL.12.10.973.
[29] V. Divya, R. Deepika, C. Yamini, P. Sobiyaa, An Efficient K-Means Clustering Initialization Using Optimization Algorithm, Proceedings of the 2019 International Conference on Advances in Computing and Communication Engineering, ICACCE 2019. (2019). https://doi.org/10.1109/ICACCE46606.2019.9079998.

Author's Profile

Mesut. Polatgil He is a lecturer in computer science in the Faculty of Applied Sciences. He received his Ph.D. degree in Numerical methods from the University of Sivas Cumhuriyet. He has more than 8 years of teaching and research experience in information systems development and machine learning. His research interests include machine learning, numerical methods and data science.

How to cite this paper: Mesut Polatgil, "Outlier Detection Algorithm Based on Fuzzy C-Means and Self-organizing Maps Clustering Methods", International Journal of Mathematical Sciences and Computing(IJMSC), Vol.8, No.3, pp. 21-29, 2022. DOI:10.5815/ijmsc.2022.03.02