Global Journal of Computer Sciences: Theory and Research

Volume 8, Issue 2, (2018) 62-69

www.gjcs.eu

Introducing a software for innovative neuro-fuzzy clustering method named NFCMR

Aref Shirazi Department of Mining and Metallurgical Engineering, Amirkabir University of Technology, Tehran 15875-4413, Iran

Adel Shirazy, School of Mining, Petroleum and Geophysics, Shahrood University of Technology, Shahrud 361999516, Iran

Shahab Saki, Department of Mining and Metallurgical Engineering, Amirkabir University of Technology, Tehran 15875-4413, Iran

Ardeshir Hezarkhani *, Department of Mining and Metallurgical Engineering, Amirkabir University of Technology, Tehran 15875-4413, Iran

Suggested Citation:

Shirazi, A., Shirazy, A., Saki, S. & Hezarkhani, A. (2018). Introducing a software for innovative neuro-fuzzy clustering method named NFCMR. *Global Journal of Computer Sciences: Theory and Research. 8*(2), 62–69.

Received from October 15, 2017; revised from February 02, 2018; accepted from July, 22, 2018.

Selection and peer review under responsibility of Prof. Dr. Doğan İbrahim, Near East University, Cyprus.

Abstract

An innovative neural-fuzzy clustering method is for predicting cluster of each new sample with the probability of its presence. This method is an attempt to first divide the samples in the region by fuzzy method with the probability of being in each cluster and then with the results of this Practice, the artificial neural network is trained, and can analyse the new data entered in the region with the probable percentage of the clusters. More clearly, after a full mineral exploration, the sample can be attributed to a certain probable percentage of anomalies. To test the accuracy of this clustering in the form of the theory alone, a case study was conducted on the results of the analysis of regional alluvial sediments data in Birjand, Iran, which resulted in satisfactory results. This software is written in MATLAB and its first application in mining engineering. This algorithm can be used in other similar applications in various sciences.

Keywords: Anomaly, clustering, NFCMR, neuro-fuzzy C-means regression.

* ADDRESS FOR CORRESPONDENCE: **Ardeshir Hezarkhani**, Department of Mining and Metallurgical Engineering, Amirkabir University of Technology, Tehran 15875-4413, Iran. ***E-mail address***: ardehez@aut.ac.ir Tel.:+98 (21) 64540

1. Introduction

One of the most important and most widely used clustering algorithms is the FCM (Fuzzy C- Means) [1] algorithm, its samples are split into a number of clusters and the number is predefined. In the fuzzy [2] version of this algorithm, the number of clusters is already specified (in this study and the proposed method, we need two clusters of anomaly and field, so the parameter C represents two clusters). Take a look at the example below to see how fuzzy clustering works. In Figure 1, a single-dimensional distribution of input samples is presented.

Figure 1. One-dimensional distribution of samples

If we use the classical *c* mean algorithm, the above data will be split into two distinct clusters, and each instance will only belong to one of the clusters. In other words, the membershipfunctionofeach samplewillhave0or1[3].Theresultoftheclassicclusteringis according to Figure 2.

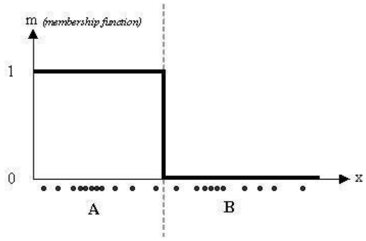

Figure 2. Classical clustering of input samples

Figure 2 shows the subordination function associated with the cluster A. The subordinate cluster b function of the complement is the membership function of A. As you can see, the input samples belong only to one of the clusters, in other words the matrix U is binary [4]. Now if we use fuzzy clustering we will have:

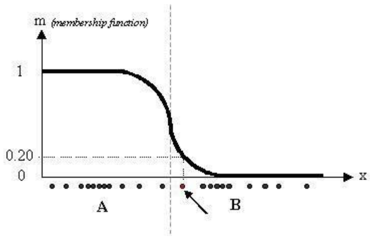

Figure 3. Fuzzy clustering of samples [5]

You can see that in this case, the curve of the membership function is smoother and the boundary between the clusters is not definitely defined. For example, a sample marked in red is attributed to the cluster B with a degree of attribution of 0.2 and to cluster A with 0.8 degree [6].

Strengths of the Fuzzy C Means Algorithm:

- Always converges.
- No monitoring algorithm.

Artificial neural networks are inspired by the animal's nervous system [7]. Artificial neural networks cover a much smaller range of biological nerve networks and have much less ability and capability than biological neural networks [8]. What is actually considered is the computational ability of the network to perform a particular activity as the approximation of a function. GRNN is often used for approximation of functions, and its structure is similar to a radial network, but in the second layer there is a small difference with the radial network [9, 10].

The difference can be seen in Figure 4 in the nprod box. In the nprod box, S2 drives in the n2 vector. Each vessel is produced by a row LW 2.1 and input vector a1 from the first layer, and all of them are normalised by the sum of the droplets a1[7, 11, 12].

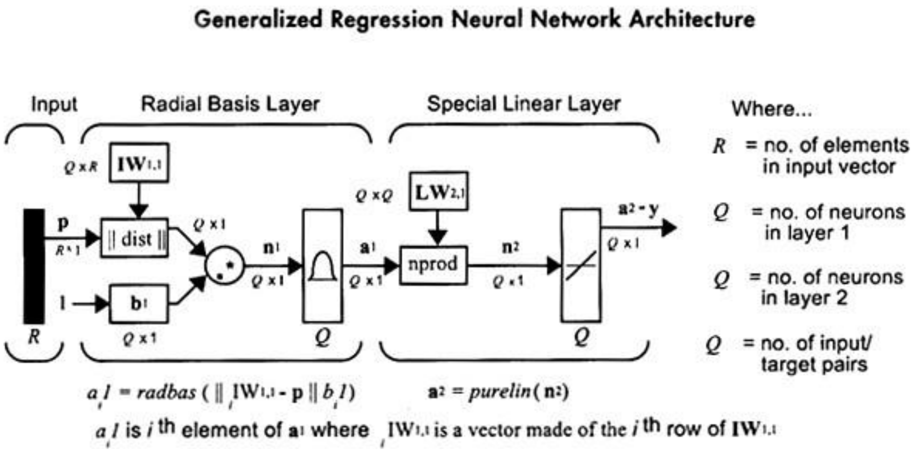

Figure 4. GRNN architecture [13]

The criteria for reaching the desired prediction are to select the appropriate radius for the network. It also provides a high degree of accuracy for training data [14].

2. Algorithm

The algorithm of work is taken as Figure 5, so that first the data of the region are divided into two groups of anomalies and fields by the FCM method and the probability percentage of the anomaly of each sample is determined, then the data and the percentage of probability. The data are then given to the GRNN [15] and then given the new sample to this system, the probability of belonging to the sample is determined by the anomaly.

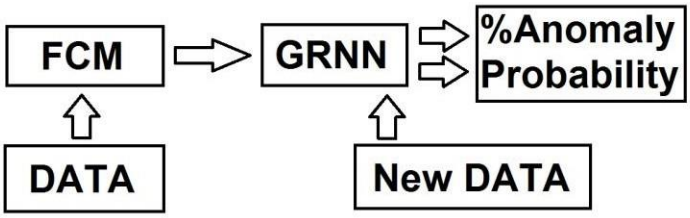

Figure 5. NFCMR flowchart

3. Location and geology of the studied area

The studied poly-metal region is located in South Khorasan province and 30 kilometers northwest of Birjand city, East Iran [16] (Figures 6 and 7).

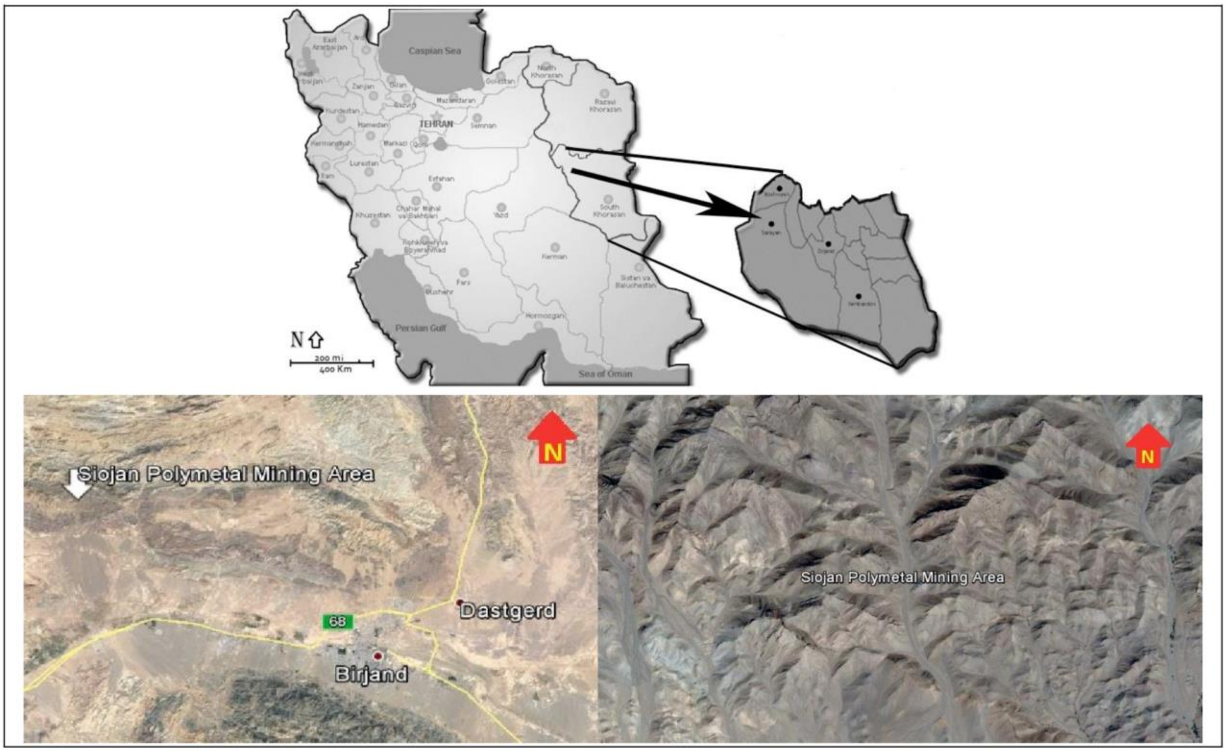

Figure 6. The studied area

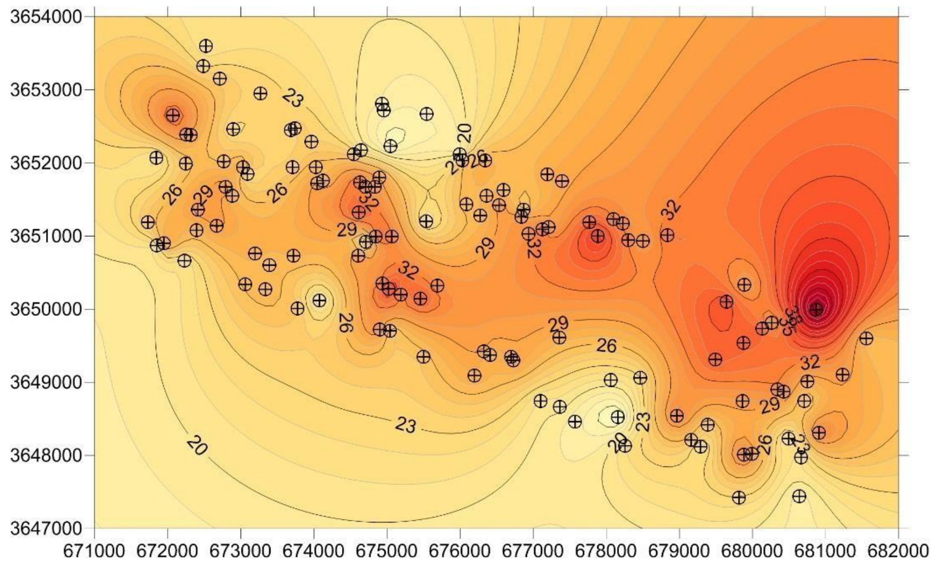

Figure 7. Place the samples taken and estimate other points by kriging

This region is located in the northwest of the Lut zone in terms of the division of the structural states of Iran's crust in the eastern part of the Central Iranian continent [17]. The Nehbandan fault system, which covers all the Sistanian states, enters the Lut zone in the northern part, with a change of direction to the west [18]. The Cheshmeh Khouri area is located in the westernmost Splay (Splay) area, a structural change that is affected by the shaft of this fault system [19].

The region's image is generally composed of volcanic and intermediate cenozoic rocks in the form of andesitic masses, basaltic andesite, dacite, which in some cases have caused alteration and mineralisation. Pyroclastic deposits such as altered tuffs and cuttings with andesite rocks are also present in the region and non-combustible units such as conglomerate, sandstone, salts and young alluvium in the region are found [20].

4. Statistical analysis

Samples were shown in logarithmic form in the statistical analysis carried out in Table 1, as well as the normality of the data.

Table 1. Primary statistical results

	Minimum	Maximum	Sum	Mean	Standard deviation	Variance	Skewness	Kurtosis
Cu	15	49.6	3272.1	27.2675	5.3913	29.066	0.45	1.49

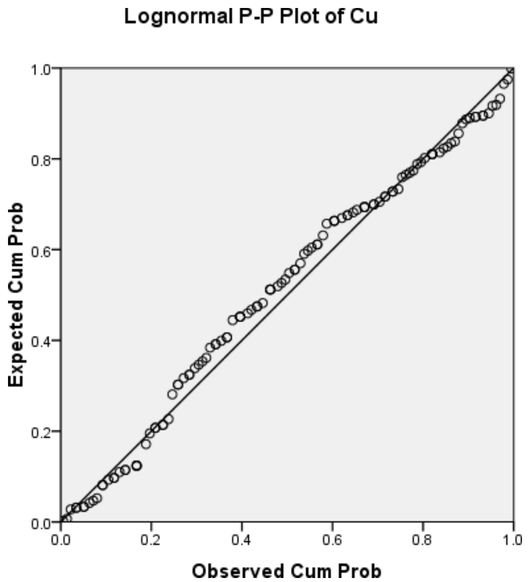

Figure 8. P–P graph as a log normal

5. Testing the NFCMR method

The data were separated as 30% as test data and 70% as training data randomly. The results were presented in Figures 9–12.

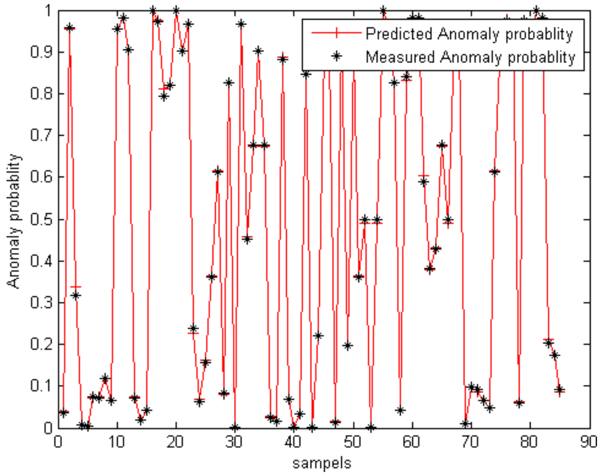

Figure 9. Forecast line and actual values (training data)

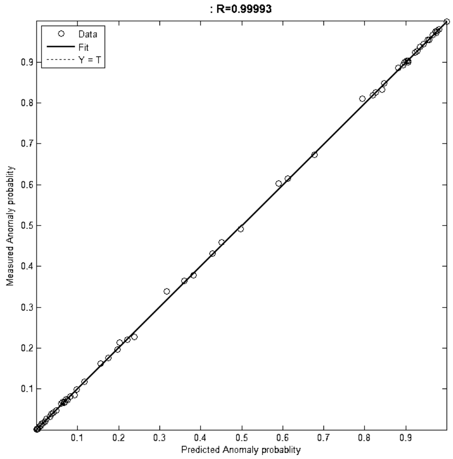

Figure 10. Predictive and real data (training data)

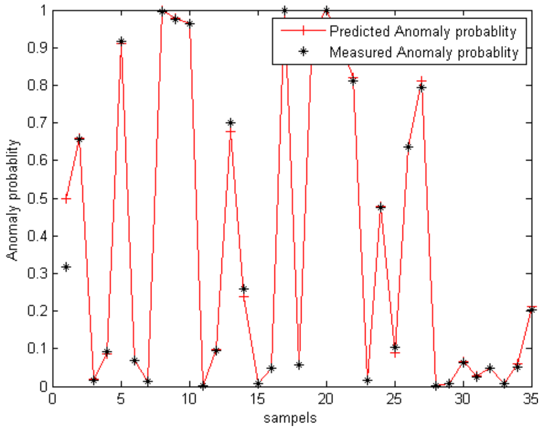

Figure 11. Forecast line and actual values (testing data)

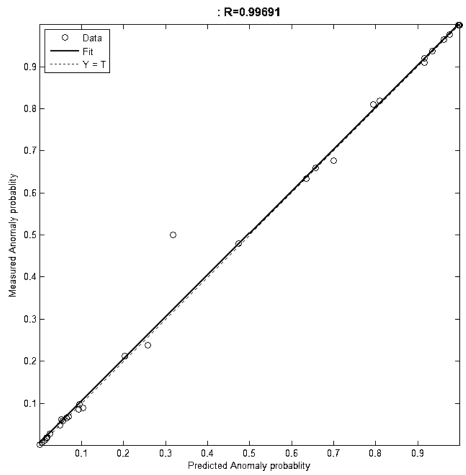

Figure 12. Predictive and real data (Testing data)

As you can see, the network has 99.99% accuracy for training data, and for test data, which is much more important than training data, it gives 99.69% accuracy, which indicates the high accuracy of this method in estimating the probability of anomalous/background sample.

6. Conclusion

The innovative method of neuro-fuzzy clustering in this paper is a method for predicting cluster (anomaly/background) of each new sample with the probability of its presence in clusters. This method, which is a combination of the Fuzzy C-Means clustering method and the General Regression Neural Network, is an attempt to first divide the samples in the region by fuzzy method with the probability of being in each cluster into two anomalies and field clusters (in Mining Engineering), and then with the results of this practice, the artificial neural network training, can analyse the new data entered in the region with the probable percentage of clusters. More clearly, after a full mineral exploration in Mining Engineering, the sample can be attributed to a certain probable percentage of anomalies. This method better demonstrates its effectiveness when it comes to extending the exploration area to other areas. However, this method can be used for any purpose and can be criticised and used with more studies for other sciences. To test the precision of this clustering as a purely theoretical one, a case study was carried out on the results of analysis of regional alluvial sediments data in Birjand, IRAN, whose accuracy was estimated as 99.99% in train data and 99.69% in test data, which yielded high accuracy results as shown.

References

[1] J. C. Bezdek *et al.*, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences,* vol. 10, issue 2–3, pp. 191–203, 1984.

[2] S. Sivanandam *et al.*, *Introduction to fuzzy logic using MATLAB* (Vol. 1). Berlin: Springer, 2007.

[3] M. Shiva *et al.*, "Fuzzy c-means cluster analysis, a robust multivariate technique in stream sediment geochemical exploration, a case study in Eastern part of Iran, Birjand," *Journal of Geology and Mining Research,* vol. 3, issue1, pp. 1–6, 2011.

[4] G. Klir and B. Yuan, (1995). *Fuzzy sets and fuzzy logic* (Vol. 4). New Jersey: Prentice hall.

[5] J. Yen and R. Langari, *Fuzzy logic: intelligence, control, and information* (Vol. 1) Upper Saddle River: Prentice Hall, 1999.

[6] L. A. Zadeh, "Fuzzy logic," *Computer*, vol. 21, issue 4, pp. 83–93, 1988.

[7] R. J. Schalkoff, *Artificial neural networks* (Vol. 1). New York: McGraw-Hill, 1997.

[8] T. M. Mitchell, “Artificial neural networks,” *Machine learning,* vol. 45, pp. 81–127, 1997.
[9] H. Demuth and M. Beale, Neural network toolbox for use with Matlab–user’s guide version 3.0, 1993.
[10] D. F. Specht, (1996). “*Probabilistic neural networks and general regression neural networks,”* paper presented at the *Fuzzy logic and neural network handbook*, 1996.
[11] A. Abraham, Artificial neural networks. *handbook of measuring system design, 2005*.
[12] M. M. Khatib *et al.*, “*Analysis of deep geological structures by variety of curie point depth in Birjand Area, East of Iran,”* Paper presented at the *6th WSEAS International Conference on geology and seismology*, Kos Island, Greece, 2012.
[13] D. F. Specht, “A general regression neural network,” *IEEE transactions on neural networks,* vol. 2, issue 6, pp. 568–576, 1991.
[14] B. Yegnanarayana, *Artificial neural networks*. Delhi: PHI Learning Pvt. Ltd, 2009.
[15] S. Saha and S. Bandyopadhyay, “A generalized automatic clustering algorithm in a multiobjective framework,” *Applied Soft Computing,* vol. 13, issue 1, pp. 89–108, 2013.
[16] J. Eftekharnezhad, 1: *100000 Geological Map of Birjand.* Tehran: Iran’s Geological Survey and Mineral Explorations Publications*,* 1986.
[17] H. Eshraghi “Auriferous sulfides from Hired gold mineralization, south Birjand, Lut block, Iran,” *Journal of Mineralogical and Petrological Sciences,* vol. 105, issue 4, pp. 167–174, 2010.
[18] R. Walker and M. Khatib, “Active faulting in the Birjand region of NE Iran,” *Tectonics,* vol. 2, issue 4, 2006.
[19] M. Emami, *Geology and petrological investigation on Shahkuh volcanic rocks, south of Birjand, eastern Iran*. Tehran: *University of Tehran, 1972*.
[20] J. Eftekhar-Nezhad and J. Stocklin, (1992). Geological map of Iran sheet K8 (Birjand), scale 1: 250,000. Tehran: Geological Survey of Iran, 1992.