Journal of Theoretical and Applied Information Technology
10th January 2015. Vol.71 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DOES EFFICIENT FUZZY KOHONEN CLUSTERING NETWORK ALGORITHM REALLY IMPROVES CLUSTERING DATA RESULT?

[1]EDY IRWANSYAH, [1]MUHAMMAD FAISAL, [2]ANNISAA PRIMADINI
[1]Dept. of Computer Science, Bina Nusantara University, Jalan KH. Syahdan No. 9 Palmerah Jakarta 11480, Indonesia
[2]Department of Electrical Engineering, University of Indonesia, Depok 16424, Indonesia
E-mail: [1]edirwan@binus.ac.id, [1]mhdfaisal93@yahoo.com, [2]primadini.annisaa@gmail.com

ABSTRACT

In this research, Fuzzy Kohonen Clustering Network (FKCN) algorithm is compared to Efficient Fuzzy Kohonen Clustering Network (EFKCN) algorithm. This research is conducted to see if EFKCN is really efficient and could do a better clustering analysis than original FKCN. We do empirical testing and simulations to compare both algorithms, by using an expanded Fisher's Iris Data. The result showed the accuracy of EFKCN is not yet better algorithm rather than FKCN.

Keywords: *Artificial Neural Network, Clustering Analysis, Fuzzy Kohonen Clustering Network, Kohonen Network, Self-orginizing Map,*

1. INTRODUCTION

Fuzzy Kohonen Clustering Network (FKCN) introduced by [1] is an unsupervised learning clustering analysis method. This algorithm integrates FCM model [2] and Kohonen network [3]. FKCN improved the error convergence as well as reduced labeling errors problem on FCM.

FKCN enhances the FCM algorithm and showed great superiority in processing ambiguity and uncertainty of a certain dataset or image. Yang et al, 2009 [4] attempted to make an efficient FKCN (EFKCN) to reduce computation process on the original FKCN. The accuracy of EFKCN achieved 92.7 percent by the fourth iteration. In which, the previous research conducted by [1] didn't show a remarkable result with the same iteration number. The improvement of learning rate computation yields a different result compared to FKCN [1]. Empirical testing with the same data set and algorithm in this research is conducted to find out whether EFKCN is really efficient and could do a better data clustering than the original FKCN.

2. RELATED WORKS

Yang et al, 2008 [4] introduced threshold values and dynamic adjustment of the weighting exponent on each fuzzy membership to adjust learning rates dynamically on FKCN by Bezdek et al, 1992 [1]. The error data of EFKCN is just 11 error data by fourth iteration. Meanwhile, FKCN has more than 40 error data with the same iteration. On the top of that, Bezdek at al, 1992 [1] had big number of mistakes in FKCN under the eighth iteration.

FKCN has been implemented well for image segmentation [5], [6], [7], [8] and [9]. Atmaca at al, 1996 [5] used FKCN to determine cluster membership values by doing some improvements, Lei at al, 1999 [6] introduced Adaptive FKCN that reduce computation process in image segmentation, by using HSV as color representation, Jabbar at al, 2009 [8] proved FKCN have ability to segment the color image and noisy color image can be segmented more effectively and provide more robust segmentation results by using FKCN with some improvements by Lu et al, 2009 [9].

de Almeida at al, 2012 [10] applied FKCN to an interval of data and proved the of FKCN is better than FCM and, Fan at al, 2013 [11] combined FKCN and motion control algorithm to develop intelligent wheelchair and Irwansyah and Hartati [12] used EFKCN to cluster the building damage hazard due to earthquake and Kriging algorithm to create building damage area zonation at Banda Aceh city, Indonesia.

Journal of Theoretical and Applied Information Technology
10th January 2015. Vol.71 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. FKCN AND EFKCN ALGORITHM

The algorithm of FKCN is summarized as follow:

Step 1: Fix c, sample space and threshold error ε > 0 some small positive constant.

Step 2: Initialize $v_0 = (v_{10}, v_{20}, ..., v_{c0})$ choose m_0 > 1 and iteration limit (t_{max})

Step 3: For t = 1,2…, t_{max}

a. Compute all learning rates:

$$m_t = m_0 - t * \Delta m, \ \Delta m = \frac{1}{t_{max}} \quad (1)$$

Where m_t is the fuzzy membership for the t iteration and Δm: fuzzy membership differences for each iteration

For each member, update its membership function:

$$u_{ik,t} = \left(\sum \left(\frac{\|X_k - V_{i,t-1}\|}{\|X_k - V_{j,t-1}\|} \right)^{\frac{2}{m-1}} \right)^{-1} \quad (2)$$

Where $u_{ik,t}$ is the membership function of the *k-th data of i-th cluster for each t-iteration*; X_k is the k-th data; $V_{i,t-1}$ is the i-th cluster center for t-1 iteration; $V_{j,t-1}$ is the j-th cluster center for t-1 iteration and m is weighting exponent on each fuzzy membership.

Compute learning rate

$$\alpha_{IK,\tau} = (\upsilon_{IK,\tau})^{\mu_\tau} \quad (3)$$

Where $\alpha_{ik,t}$ is the learning rate of the *k-th data* of *i*-th cluster for t iteration.

b. Update all weight vectors:

$$v_{i,t} = v_{i,t-1} + \frac{\sum_{k-1}^{n} \alpha_{ik,t}(x_k - v_{i,t-1})}{\sum_{j=1}^{n} \alpha_{ij,t}} \quad (4)$$

Where v_i is i-th cluster center and $v_{i,t}$ is i-th cluster center for t iteration

c. Compute the function

$$E_t = \|V_t - V_{t-1}\| \quad (5)$$

Where E_t is error for each t iteration; V_t is cluster center for t iteration and V_{t-1} and cluster center for t-1 iteration.

d. If $E_t < \varepsilon$ stop. Else t = t + 1 goto step 3

The difference between EFKCN and FKCN lies in the learning rate computation, while FKCN uses equation (3), EFKCN follows the equation (6) below:

$$\alpha_{IK,\tau} = \begin{cases} (u_{ik,t})^{m_u} & u_{ik,t} > t_d \\ (u_{ik,t})^{m_t} & t_d \le u_{ik,t} \le t_u \\ (u_{ik,t})^{m_d} & u_{ik,t} < t_d \end{cases} \quad (6)$$

Where t_u and t_d are threshold value, t_u is upper cut set (0.5,1), t_d is lower cut set (0,0.5), m_u and m_d are fuzzy convergence operators, m_u (0,1) and m_d > m_t

4. EXPERIMENTS AND RESULT

Fisher's IRIS data [13] is used for the simulations and testing. The IRIS will be expanded 40 times to become a set of 6000 sample data. Simulations of data are executed by using MATLAB R2013a on Intel® Core™ i3 CPU 2.27 Ghz PC with Windows 7 Professional 64-bit operating system and memory of 3072 MB RAM. The result will be compared to the actual cluster center by Hathway and Bezdek, 1995 [14].

Journal of Theoretical and Applied Information Technology
10th January 2015. Vol.71 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1. Result comparison to Yang et al, 2009

	FKCN Yang et al (2009)	FKCN Irwansyah et al (2014)	EFKCN Yang et al (2009)	EFKCN Irwansyah et al (2014)
Parameter	m=3; ε=0.001; tmax=50; c=3	m=3; ε=0.001; tmax=50; c=3	m=3;ε=0.001; tmax=50; c=3; tu=0.7; td=0.3; mu=0.4;	m=3; ε=0.001; tmax=50; c=3; tu=0.7; td=0.3; mu=0.4;
Correct Number	5480	5427	5560	4513
Correct Rate	91.33%	90.45%	92.67%	75.22%
Iteration Number	12	12	4	4
Hathway Cluster Center	5.00 5.93 6.58 3.42 2.77 2.97 1.46 4.26 5.55 0.24 1.32 2.02			
Cluster Center	5.01 5.75 6.60 3.14 2.73 3.01 1.48 4.14 5.36 0.25 1.29 1.91	5.04 5.91 6.58 3.38 2.78 2.99 1.60 4.37 5.36 0.30 1.40 1.91	5.00 5.91 6.58 3.41 2.79 3.00 1.48 4.22 5.44 0.25 1.32 1.98	5.32 5.88 6.14 3.24 3.00 2.97 2.38 3.92 4.51 0.63 1.26 1.51
The square sum of the relative cluster center's error	0.1073	0.1208	0.0256	0.0787

Table 2. The clustering result comparison with same iteration number

	FKCN	EFKCN
Parameter	m=3; ε=0.001; tmax=50; c=3	m=3;ε=0.001; tmax=50; c=3; tu=0.7; td=0.3; mu=0.4;
Correct Number	5456	5202
Correct Rate	90.93%	86.70%
Iteration Number	10	10
Hathway Cluster Center	5.00 5.93 6.58 3.42 2.77 2.97 1.46 4.26 5.55 0.24 1.32 2.02	
Cluster Center	5.04 5.97 6.61 3.37 2.79 3.00 1.60 4.45 5.41 0.30 1.44 1.94	5.01 6.02 6.57 3.37 2.79 2.97 1.58 4.55 5.37 0.28 1.49 1.92
The square sum of the relative cluster center’s error	0.0975	0.0135

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

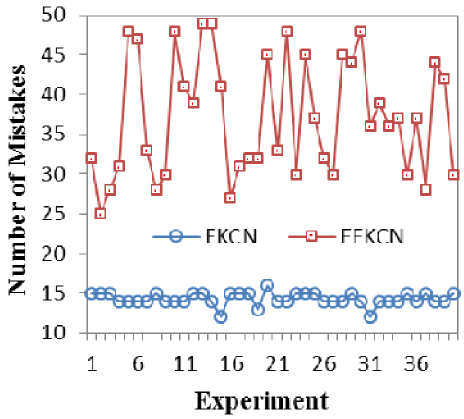

Figure 1: Number of mistakes FKCN with t = 12 and EFKCN with t = 4

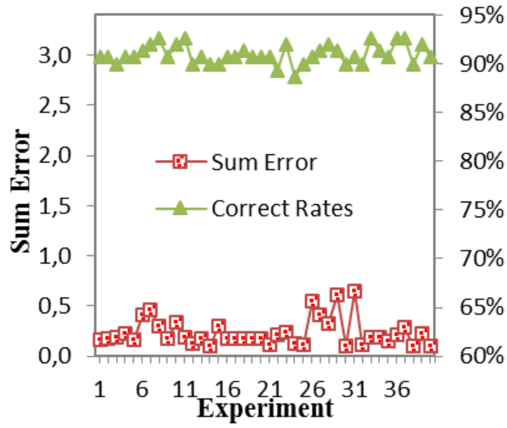

Figure 3: Comparison between square sum error and correct rates using FKCN

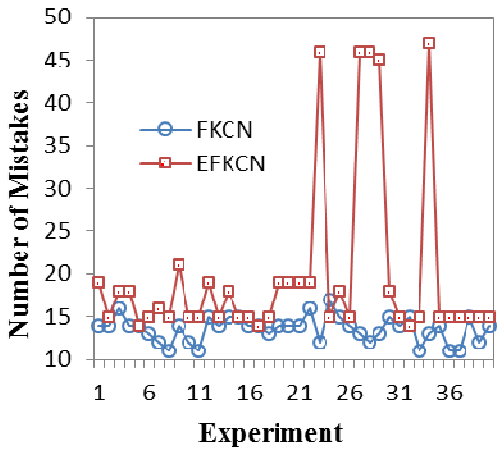

Figure 2: Number of mistakes FKCN and EFKCN with t = 10

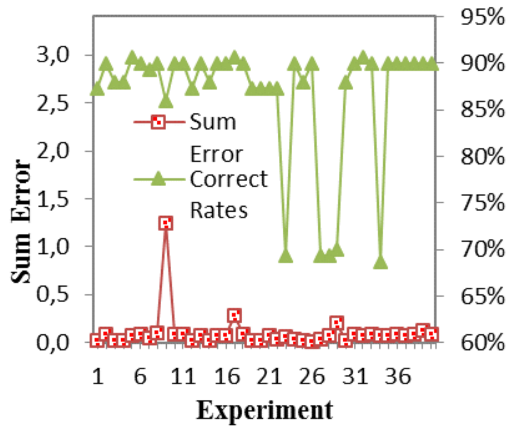

Figure 4: Comparison between square sum error and correct rates using EFKCN

The experimental results show the accuracy of EFKCN [4] not yet close to 92.7 percent by fourth iteration (Table 1). Moreover, with t = 4, the accuracy of EFKCN is 75.22 percent. Meanwhile, the accuracy of FKCN [1] with t = 12 achieved 90.45 percent; just slightly different from result by Yang et al, 2009 [4]. Figure 1 showed that number of mistakes in FKCN is always smaller than EFKCN and the square sum error of the relative cluster centers are the smallest number from the experimental result (See Table 1 and Table 2).

Table 2 described the result of simulation and testing with same iteration (t = 10). Once again, the result showed the accuracy of FKCN is better than EFKCN, 90.93 percent compared with 86.70 percent. Interestingly, eventhough EFKCN has a lower correct rate than FKCN, computation results showed EFKCN square sum of the relative cluster center is very low.

5. CONCLUSION

From the experimental results with same data, either having a different or the same iteration number, EFKCN by Yang et al, 2008 didn't really improve of the accuracy compared to FKCN by Bezdek et al, 1992. Threshold value and fuzzy convergence operators that are proposed by EFKCN indeed produced small square sum error but couldn't reach higher correct rates than FKCN. Hence, EFKCN is not yet a better algorithm than FKCN.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFRENCES:

[1] Bezdek, J.C., Tsao, E.C-K., Pal, N.R.: Fuzzy Kohonen Clustering Networks. IEEE (1992) 1035-1043

[2] Bezdek, J. C.: A Convergence Theorem for the Fuzzy C-Means Clustering Algorithm. IEEE (1980)

[3] Kohonen, T.: Self-Organization and Associative Memory. 3rd Edition. Springer, Berlin (1989)

[4] Yang, Y., Jia, Z., Chang, C., Qin, X., Li, T., Wang, H., Zhao, J.: An Efficient Fuzzy Kohonen Clustering Network Algorithm. IEEE (2008) 510 - 513

[5] Atmaca, H., Bulut, M., Demir, D.: Histogram Based Fuzzy Kohonen Clustering Network For Image Segmentation. IEEE (1996) 951-954

[6] Lei, W., Feihu, Q.: Adaptive Fuzzy Kohonen Clustering network for Image Segmentation. IEEE (1999) 2664-2667

[7] Xiaodan, W., Hua, J., Rongchun, Z.: Texture Segmentation method Based on Incomplete Tree Structured WaveleTransform and Fuzzy Kohonen Clustering Network. IEEE (2000) 2684-2687

[8] Jabbar, N., Ahson, S.I., Mehrota, M.: Fuzzy Kohonen Clustering Network for Color Segmentation. IPCSIT vol. 3. LACSIT Pres (2011) 254-257

[9] Lu, B., Wei, Y., Li, J.: A Noise-Resistant Fuzzy Kohonen Clustering Network Algorithm for Color Image Segmentation. IEEE (2009) 44-48

[10] de Almeida, C.W.D., Souza, R.M.C.R., Candeias, A.L.B.: IFKCN: Applying Fuzzy Kohonen Clustering Network to Interval Data. IEEE (2012)

[11] Fan, J., Jia, S., Li, X.: The Application of Fuzzy Kohonen Clustering Network for Intelligent Wheelchair Motion Control. IEEE (2013) 1995-2000

[12] Irwanyah, E., Hartati, S.: Data Clustering and Zonation of Earthquake Building Damage Hazard Area Using FKCN and Kriging Algorithm. In: M.K. Kundu et al. (eds.) Advanced Computing, Networking and Informatics - Volume 1, Smart Innovation, Systems and Technologies 27. Springer International, Switzerland (2014) 171-178.

[13] Fisher, R.A.: The Use of Multiple Measurements in Taxonomic Problems. Annual Eugenics, 7, Part II, (1936) 179-188

[14] Hathway, J. R., Bezdek, J. C.: Optimation of Clustering Criteria by Reformulation. IEEE (1995) 241 – 245