Integrating type-1 fuzzy and type-2 fuzzy clustering with k-means for pre-processing input data in classification algorithms

Vahid Nouri[1, *], Mohammad Reza Akbarzadeh[2], Tootoonchi, Alireza Rowhanimanesh[3]

[1] Department of Computer Engineering, Islamic Azad University, Mashhad Branch, Mashhad, Iran
[2] Department of Electrical Engineering, University of Neyshabur, Neyshabur, Iran
[3] Departments of Electrical and Computer Engineering, Center of Excellence on Soft Computing and Intelligent Information Processing (SCIIP), Ferdowsi University of Mashhad, Mashhad, Iran

Email address:

vahid.nouri@mshdiau.ac.ir (V. Nouri), akbarzadeh@ieee.org (M. R. Akbarzadeh), rowhanimanesh@ieee.org (A. Rowhanimanesh)

Abstract: In several papers, clustering has been used for preprocessing datasets before applying classification algorithms in order to enhance classification results. A strong clustered dataset as input to classification algorithms can significantly improve the computation time. This can be particularly useful in "Big Data" where computation time is equally or more important than accuracy. However, there is a trade-off between computation time (speed) and accuracy among clustering algorithms. Specifically, general type-2 fuzzy c-means (GT2 FCM) is considered to be a highly accurate clustering approach, but it is computationally intensive. To improve its computation time we propose a hybrid clustering algorithm called KFGT2FCM that combines GT2 FCM with two fast algorithms k-means and Fuzzy C-means algorithm for input data preprocessing of classification algorithms. The proposed algorithm shows improved computation time when compared with GT2 FCM on five benchmarks from university of California Irvine (UCI) library.

Keywords: Classification, Input Data Preprocessing, Clustering, General Type-2 Fuzzy Logic, Fuzzy C-Means (FCM), K-Means

1. Introduction

Classification is a common problem in data mining [20] where datasets are mapped into predefined groups called classes. Classes are defined according to the similarity of characteristics or features of data [15]. Since the classes are determined before applying the real data, this method is known as a supervised learning algorithm. Classification is used in many fields and sciences such as, image segmentation [1, 2], geology [3], robot control [4, 5], bio-informatics [6], genetics [8], biology [7] and healthcare [9].

Several researches have shown that the speed improvement of a classification algorithm is enhanced if the input data is first clustered before classification. This is particularly applicable when handling big data, where low computation time is equally or more important than classification accuracy.

The class information also improves the accuracy of clustering [10]. In order to have the advantages of clustering and classification, many hybrid algorithms have been developed [10]. For example in both [11] and [10], first the criterion is preprocessed and optimized by a clustering algorithm and then in the next step the classification criterion is applied to the achieved clustering results to enhance the accuracy of classification algorithms.

Generally, accuracy and computation time of clustering algorithms are in contrast with each other, i.e. the higher the accuracy, the more computation time. Two well-known clustering algorithms are k-means and FCM. FCM and k-means are fast but have low accuracy [15]. General type-2 fuzzy clustering (GT2 FCM) is a new method that has high accuracy but is computationally intensive. In [18], a general type-2 fuzzy clustering algorithm is introduced that is based on α-planes. This algorithm has high accuracy and can deal

with the uncertainty in datasets, while k-means and FCM, which are fast clustering algorithms, cannot handle the uncertainty in a dataset.

There are several works that concentrate on improving the computation time of type-2 fuzzy clustering. A modified version of type-2 fuzzy system was proposed in [21] to improve the speed (computation time) of type-2 fuzzy clustering. Also, in other studies [12, 13, and 14] interval type-2 fuzzy is used instead of general type-2 fuzzy for clustering, because interval type-2 is faster than general type-2 fuzzy.

In addition Yang worked on similarity metrics of type-2 fuzzy clustering algorithms on fuzzy datasets [23-26]. In these studies, Yang redefined new similarity metrics based on union maximum. These new similarity metrics affect type-2 fuzzy clustering efficiency.

The hybrid clustering algorithm which is proposed in this paper is used for data input preprocessing of classification algorithms to address the high computation time of general type-2 fuzzy clustering algorithm. The proposed hybrid method is based on a combination of general type-2 fuzzy, which is an accurate algorithm and k-means, which is a fast algorithm. We call the proposed approach KFGT2FCM. KFGT2FCM has the advantages of general type-2 fuzzy, k-means and FCM clustering algorithms, i.e. it has high accuracy and low computation time. The results are compared with GT2 FCM clustering algorithms for different datasets. Unlabeled datasets are used for clustering algorithms, however, labeled datasets are used for classification algorithms. While, we use classification datasets in our experiments, we can measure the accuracy of our clustering algorithm. The paper is organized as follows: section 2 discusses the proposed hybrid algorithm. The results and conclusion are presented in sections 3 and 4, respectively.

2. Proposed Method

Our method is based on k-means, Fuzzy C-means and general type-2 fuzzy clustering. General type-2 fuzzy clustering was presented in [18]. First, a general overview of type-2 fuzzy is presented, and then the proposed method is described.

There are two kinds of type-2 fuzzy sets which are used in clustering algorithms: 1) interval and 2) general. In interval type-2 fuzzy, the secondary membership function always equals one, while in general type-2 fuzzy it is a value in the interval of [0,1].

General type-2 fuzzy clustering is based on FCM (Fuzzy C-Means) algorithm. Like FCM, it initializes the centers randomly. The FCM algorithm uses linguistic terms such as "Small", "Medium" or "High", modeled by type-1 fuzzy sets for the fuzzifier parameter M (Figure 1). The FCM algorithm is used by the GT2 FCM cluster membership functions. The general type-2 fuzzy clustering proposed in [18] uses α-planes. The uncertainty of general type-2 fuzzy sets is managed by α-planes. The GT2 FCM algorithm exploits the linguistic fuzzifier M for its secondary membership functions of the general type-2 fuzzy partition matrix $\tilde{u}_j$ as shown in Equation 1. In addition, Equation 2, that is a membership grade $\tilde{u}_j(x_i)$ is expressed as type-1 fuzzy sets, which is used to describe the membership degree of pattern x_i to cluster v_j.

$$\tilde{u}_j = \sum_{x_i \in X} \tilde{u}_j(x_i) \tag{1}$$

$$\tilde{u}_j(x_i) = \bigcup_{\alpha \in [0,1]} {}^{\alpha}/_{S_{\tilde{u}_j}(x_i|\alpha)} = \bigcup_{\alpha \in [0,1]} {}^{\alpha}/_{[S_{\tilde{u}_j}^L(x_i|\alpha), S_{\tilde{u}_j}^R(x_i|\alpha)]} \tag{2}$$

Where $s_{\tilde{u}_j}^R(x_i \mid \alpha)$ and $s_{\tilde{u}_j}^L(x_i \mid \alpha)$ are calculated by (3) and (4):

$$s_{\tilde{u}_j}^R(x_i|\alpha) = max\left(\frac{1}{\sum_{l=1}^{c}(d_{ij}/d_{il})^{2/(s_M^L(\alpha)-1)}}, \frac{1}{\sum_{l=1}^{c}(d_{ij}/d_{il})^{2/(s_M^R(\alpha)-1)}}\right) \tag{3}$$

$$s_{\tilde{u}_j}^L(x_i|\alpha) = min\left(\frac{1}{\sum_{l=1}^{c}(d_{ij}/d_{il})^{2/(s_M^L(\alpha)-1)}}, \frac{1}{\sum_{l=1}^{c}(d_{ij}/d_{il})^{2/(s_M^R(\alpha)-1)}}\right) \tag{4}$$

Based on [18], centroid $C_{\tilde{u}_j}$ can be calculated as a weighted composition of the interval centroids of individual
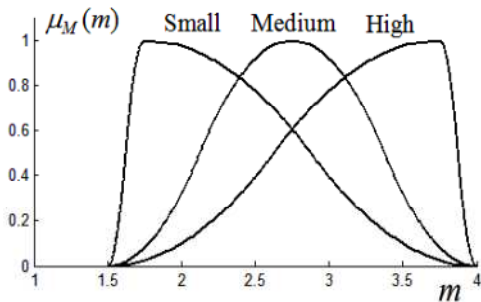

Figure 1. Linguistic variables for initializing the membership functions [18]

α-planes using Equation 5. The input of Equation 5 is $\tilde{u}_j$. Here, d_{ij} is the distance of ith data from jth centroid. Initial centroids are used for the first iteration. s_M^R and s_M^L are obtained as shown in Figure 2 for each α-planes and c is the number of clusters.

To compute the precise cluster position, Equation 6 is used to defuzzify the cluster centroid $C_{\tilde{u}_j}$.

$$C_{\tilde{u}_j} = \bigcup_{\alpha \in [0,1]} {}^{\alpha}/_{[c_{\tilde{u}_j}^L(\alpha), c_{\tilde{u}_j}^R(\alpha)]} \tag{5}$$

$$v_j = \frac{\sum_{i=1}^{K} y_i C_{\tilde{u}j}(y_i)}{\sum_{i=1}^{K} C_{\tilde{u}j}(y_i)} \qquad (6)$$

In Equation 6, K shows the number of steps that the domain of the centroid has been discretized into and y_i is the position vector of i_{th} discretized step. According to [18], the hard-partitioning is done based on the defuzzified value of the type-1 fuzzy membership grade. So, the following rule is used for hard-partitioning:

$$\text{If } \left(\tilde{u}_j(x_i)\right) > \left(\tilde{u}_k(x_i)\right))), \ k=1,\dots,c, \ k\neq j \qquad (7),$$

Then belongs to cluster j

The authors, in [18] use Equation 8 for hard-partitioning instead of Equation 7. In Equation 7, since the Euclidian distance norm is used to calculate the membership of pattern x_i to cluster j in the multidimensional space, it seems redundant to separately aggregate identical membership values for each dimension.

Therefore, in [18] the authors use Equation 8 for hard-partitioning:

$$\text{If } (c\left(\tilde{u}_j(x_i)\right) > c\left(\tilde{u}_k(x_i)\right))), \ k=1,\dots,c, \ k\neq j \qquad (8)$$

Then belongs to cluster j

The centroid of the type-1 fuzzy membership grade $c(\tilde{u}_j(x_i))$ can be calculated using Equation 9:

$$c\left(\tilde{u}_j(x_i)\right) = \frac{\sum_{i=1}^{K} y_i \tilde{u}_j(y_i)}{\sum_{i=1}^{K} \tilde{u}_j(y_i)} \qquad (9)$$

In this equation, K and y_i have the same definitions as in Equation 6, where $c(\tilde{u}_j(x_i))$ is the centroid of the jth cluster.
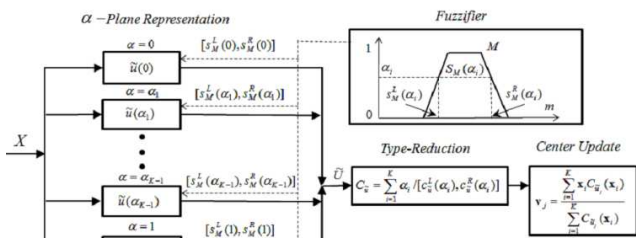

Figure 2. *Schematic view of GT2 FCM [18]*

Schematic view of GT2 FCM is depicted in Figure 2 Because of initial centroids are selected randomly the algorithm has more number of iterations, hence, more computation time (lower speed). So, if a clustering algorithm, such as FCM, finds the centers one step before GT2 FCM and passes them to GT2 FCM, the computational time of GT2 FCM will be reduced. But, FCM clustering algorithm has higher computational time against a simple clustering

algorithm same as k-means. Therefore, in order to make FCM faster than standard FCM, a k-means clustering algorithm is run before FCM and finds the centers. K-means is used to determine the centroids of input data and then calculate the distances of each data from all centroids. The normalized distances are assumed as initial values of membership functions of input data of FCM. By doing so, FCM would have a better starting point and it helps to reduce the execution time and iterations of FCM. This algorithm is called KFGT2FCM.

k-means is one of the most common algorithms in clustering. In this method, k denotes the number of clusters. k-means algorithm has three steps including:

Step 1) k cluster centers are specified, randomly i.e. one center for each cluster, step 2) for each input, and distance from each cluster center is calculated. The data belongs to the cluster which has the minimum distance from the center. This step is repeated for all dataset, and step 3) the barycenters of clusters (which are generated in step 2) are calculated and considered as new cluster centers and then the algorithm goes to step 2 [15]. These steps are repeated until centers do not change for the two consecutive iterations. The algorithm minimizes its cost function to achieve the target. Its cost function denoted as Equation 10 [15].

$$J = \sum_{i=1}^{n} \sum_{j=1}^{K} \left\| x_i - c_j \right\|^2 \qquad (10)$$

Here, n shows the number of samples, K is the number of clusters, c_j shows the jth cluster and x_i is ith sample of pattern.

In this paper we use Euclidian distance which is a traditional metric for distance measurement of k-means. Euclidian distance is presented in Equation (11) [16]:

$$d(x,y) = \sqrt{\sum (x_i - y_i)^2} \qquad (11)$$

D_{ij} is the Euclidian distance between ith sample with jth cluster.

Also, in FCM clustering algorithm the cluster numbers (c) are determined before clustering, same as K-means. The objective function of FCM is denoted as Equation 12:

$$J_0 = \sum_{i=1}^{d} \sum_{j=1}^{c} \mu_{ij}^m \left\| x_i - v_j \right\|^2 \| \qquad (12)$$

In Equation 12, m is a real number which is greater than one and by default is equal to two.

X_i is the ith sample and V_j is the jth cluster's centroid. Parameter of μ_{ij} shows the membership function of ith sample of jth cluster. The sign of $\|*\|$ denotes the similarity of the sample with the cluster centroid which can be implemented using different functions. The similarity function that is used in this paper is the Euclidean distance. The μ_{ij} are elements of a two dimensional matrix called μ. The μ_{ij} values are numbers between zero and one [6, 9]. To calculate the value of μ and centroids of clusters, the Equations 13 and 14

are used, consecutively.

$$\mu_{ij} = \left[\sum_{k=1}^{c} \left(\frac{\|x_i - v_j\|^2}{\|x_i - v_k\|^2} \right)^{1/(m-1)} \right]^{-1} \tag{13}$$

$$vj = \frac{\sum_{i=1}^{d} u_{ik}^{m} x_i}{\sum_{i=1}^{d} u_{ij}^{m}} \tag{14}$$

In GT2 FCM, the centroids are initialized using random values, similar to k-means and FCM. In the proposed method, first, k-means is applied to the dataset using Euclidean distance to cluster the input dataset. We use K-means instead of FCM due to its higher speed, however it has lower accuracy against FCM. Therefore, with higher speed and lower accuracy the center of clusters are detected. The obtained centers of k-means are applied to the FCM for initializing the centroids, but, in FCM, first the membership function degree of dataset must be initialized. Usually, in the standard FCM the membership function degrees are initialized using random values.
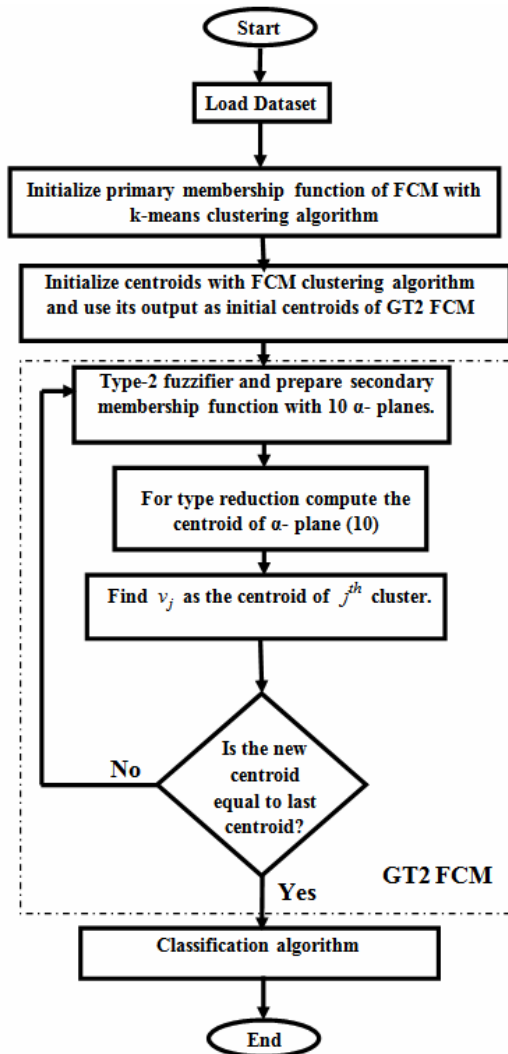

Figure 3. *Flowchart of proposed KFGT2FCM*

In the proposed method, after using k-means at the beginning of algorithm, the distance of each data to each cluster is calculated, normalized and used as initial centroids of FCM. Therefore, the FCM clustering becomes faster. In the next step, the obtained centroids of FCM are used as initial centroids of GT2 FCM. Then, the type-2 fuzzifier function calculates the secondary membership functions based on α-planes and using Equation 3, Equation 4 and "Medium" linguistic term for secondary membership function as depicted in Figure 1. We use 10 α-planes. Furthermore, EKM [1] algorithm [22] is used for type reduction and finding the centroids of α-planes. EKM was introduced by Mendel and Wu to enhance the computation time of KM. EKM is 39% faster than KM algorithm and saves about two iterations while KM find the answer usually between two to six iterations [22]. In this paper, Equation 5 which is based on EKM algorithm is used for type reduction to find the centroids of 10 α-planes. In this way, the type-2 fuzzy membership function reduces to a primary membership function which is type-1 fuzzy. To find the precise center of each cluster, the centroids should be determined using Equation 6. The centroids calculated by Equation 6 are compared to the previous centroids of each cluster. If they are not equal, the algorithm recalculates the secondary membership function using new centroids. Then the previous steps are repeated. Otherwise, the algorithm finishes (Figure 3).

The centroids which are applied to GT2 FCM are closer to optimal centroids compared to random centroids. Therefore, the required time for GT2 FCM to find the optimal centroids is decreased. Actually, k-means makes FCM faster and FCM makes GT2 FCM faster, consecutively. The flowchart of KFGT2FCM is depicted in Figure3.

3. Simulation Results

In this section the experimental setup and simulation results are presented.

3.1. Experimental Setup

Table 1. *List of datasets that used for experiments*

Dataset	Attributes	Size	Clusters
Iris	4	150	3
Wine	13	178	3
Pima Indian	8	768	2
Magic	10	19,020	2
Shuttle	9	43,500	2

In this paper, five standard datasets of university of California Irvine (UCI) are selected, including Iris, Wine, Pima Indians, Shuttle and Magic which have been listed in table 1 [19]. The Shuttle data has been divided into two classes. One class (class 1) includes the most numerous data class which is 80% of data and the second class (class 2) contains the remaining less numerous data classes which is the 20% of data. All of the datasets of Table 1 are applied to GT2 FCM,

---

[1] Enhancement Karnik Mendel

KFGT2FCM 50 times. The system used for doing the experiments and simulations is an Acer 5750G system with an Intel Core i7-2630QM@2.00GHz and 6.00 GB RAM and running Windows 7. MATLAB software has been used for implementing the algorithms. For fair comparisons of computation time of the two algorithms, the target accuracy has been assumed the same for both algorithms (i.e. GT2 FCM, KFGT2FCM) are based on GT2 FCM [18], and use the same membership functions. Since initial centroids of k-means and GT2 FCM are selected randomly, we run each algorithm for 50 iterations, i.e. with 50 sets of random initial centroids, to show that the random initial centroids have trivial effects on the results. For our experiments we use 30% and 70% of each dataset. In this way, the effect of number of samples is observed better.
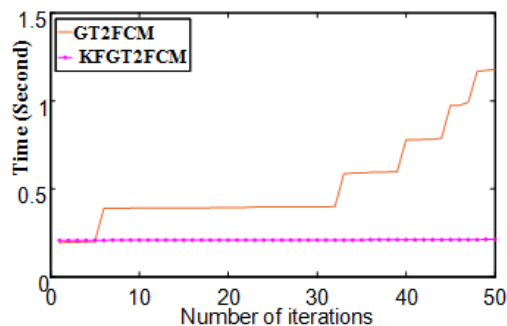
3.2. Experimental Results

Figure 4. *Comparison of computation time for 50 iterations of both algorithms with a target accuracy of 60% for 230 data of Pima Indians dataset*
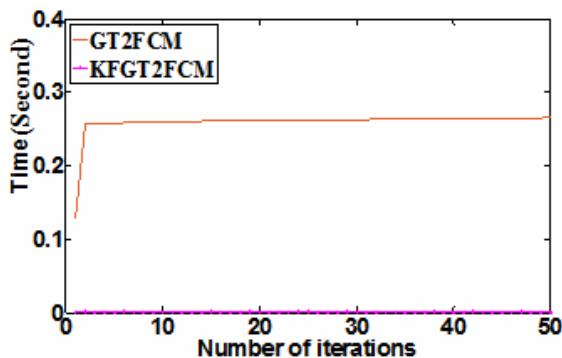

Figure 5. *Comparison of computation time for 50 iterations of both algorithms with a target accuracy of 66% for 53 data of Wine dataset*

The computation time of two algorithms are shown in Figure 4, Figure 5, Figure 6 and Figure 7. The 30% of Pima Indians dataset which have been selected randomly is applied to the two algorithms while the target accuracy is assumed to be 60% for both. In Figure 4 the results of this experiment is depicted. In another experiment which is shown in Figure 5,

30% of Wine dataset which have been selected randomly is applied to the both algorithms while the target accuracy is assumed to be 66% for two algorithms.

Comparing Table 2 and Table 3, reveals that KFGT2FCM outperforms GT2 FCM significantly for low target accuracies. For the experiments performed for generating results of Table 2 and 3, 30% and 70% of each dataset which selected randomly were used, respectively. However, for the experiments done for generating results of Figure 4, Figure 5, Figure 6 and Figure 7, 30% or 70% of each dataset, which selected randomly, were exploited.

As both pictures illustrate the computation time of KFGT2FCM is less than GT2 FCM. Also, in Figure 6 and Figure 7, 70% of Iris and 30% of Wine dataset are applied to both algorithms, respectively. The comparison shows that proposed method is faster that GT2 FCM.

Tables 2 and 3 show that for 60% of cases the proposed method (KFGT2FCM) is faster than GT2 FCM. Also, for 40% of cases where KFGT2FCM is slower for them than GT2FCM, the computation time difference of these algorithms is little.

In Table 2 and 3 Acc shows the target accuracy for each dataset in 50 iterations. Also, each the best results in each table are bolded. All of the results are calculated in seconds.
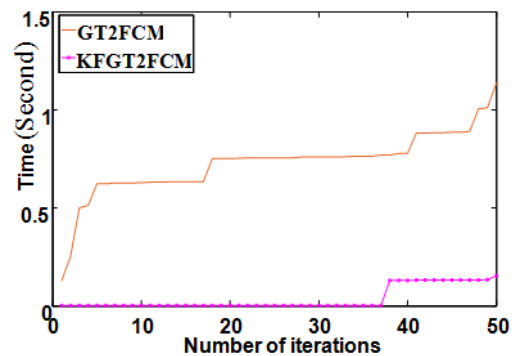

Figure 6. *Comparison of computation time for 50 iterations of both algorithms with a target accuracy of 75% for 105 data of Iris dataset*
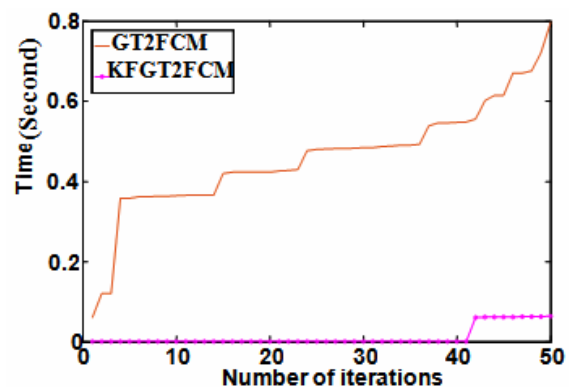

Figure 7. *Comparison of computation time for 50 iterations of both algorithms with a target accuracy of 80% for 45 data of Wine dataset*

Table 2. *Comparing computation time (in seconds) of two algorithms on 30% of datasets*

	Iris Acc:70%	Wine Acc:60%	Pima Indians Acc: 62.5%	Shuttle Acc: 55%	Magic Acc: 50%
GT2 FCM	0.1459	0.26	3.1795	2.32e-4	2.13e-5
KFGT2FCM	1.61e-5	1.55e-5	0.8244	2.8e-5	2.32e-5

*Table 3.*Comparing computation time (in seconds) of two algorithms on 70% of datasets

	Iris Acc:70%	Wine Acc:60%	Pima Indians Acc: 62.5%	Shuttle Acc: 55%	Magic Acc: 50%
GT2 FCM	1.1	0.53	1.63e-5	2.32e-4	1.29e-4
KFGT2FCM	0.04	1.64e-5	1.75e-5	2.33e-4	1.37e-4

4. Discussion and Conclusion

Recently, several works have used clustering and classification in sequential structures to improve the efficiency of classification algorithms. As indicated, the performance of classification learning is enhanced if the input data is first clustered and then used for classification. However, there is a trade-off between computation time (speed) and accuracy of clustering algorithms. In this paper, a new clustering method is introduced to improve the computation time of a classification algorithm by preprocessing classification dataset. To address the conflict of high computation time and high accuracy of clustering algorithm, we propose a hybrid clustering algorithm called KFGT2FCM.

This hybrid algorithm is a combination of high accuracy general type-2 fuzzy C-means (GT2 FCM) that can handle the uncertainty via using α-planes with low computation time k-means algorithm for input data preprocessing of classification algorithms. The proposed algorithm improves the speed of GT2 FCM. It has been evaluated using five datasets of UCI for clustering with different target accuracy.

For cases were 30% and 70% of data of each dataset are used, KFGT2FCM obtains better results compared to the GT2 FCM when target accuracy is low or features size are small. Also, results show that the number of clusters affects computation time of both algorithms.

The results depict that the proposed method is significantly faster than GT2 FCM, also in the 70% of case studies KFGT2FCM is faster than both GT2 FCM and in remaining 30% of case studies the difference of computation time of KFGT2FCM is not very high. According to Table 1, KFGT2FCM is more suitable for the most kind of low accuracy target of datasets than GT2 FCM.

Acknowledgements

References

[1]   Ting liu, Jurrus, E., Seyedhosseini, M., Ellisman, M., Tasdizen, T.,Watershed merge tree classification for electron microscopy image segmentation. 21st International Conference on Pattern Recognition (ICPR), 11-15 Nov. 2012.

[2]   Zhang Bin,Ma, Guorui,Zhang, Zhi,Qin, Qianqing, Region-based classification by combining MS segmentation and MRF for POLSAR images , Journal of Systems Engineering and Electronics. 2013 (Volume:24, Issue:3).

[3]   Zhou Q. ,Tong G. ,Xie D. ,Li B. , A Seismic-Based Feature Extraction Algorithm for Robust Ground Target Classification . Signal Processing Letters, IEEE. 2012 (Volume: 19, Issue: 10).

[4]   Garcia Bermudez, F.L., Julian, R.C., Haldane, D.W., Abbeel P., Performance analysis and terrain classification for a legged robot over rough terrain , International Conference on Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ.7-12. 2012, Vilamoura, Portugal.

[5]   Nasrollahi P., Jafari S. , Ebrahimi, M., Action classification of humanoid soccer robots using machine learning , 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP), 2-3 May 2012, Shiraz, Iran.

[6]   B., H.B., N., J.C.,   Hierarchical classification using a Competitive Neural Network , Eighth International Conference on Natural Computation (ICNC), 2012. 29-31 May 2012, Chongqing ,China.

[7]   W. Yang , K. Wang , W. Zuo, Prediction of protein secondary structure using large margin nearest neighbor classification . Advanced Computer Control (ICACC), 2011 3rd International Conference, 18-20 Jan. 2011, Harbin, China.

[8]   Yuvaraj, N. ,Vivekanandan, P.,   An efficient SVM based tumor classification with symmetry Non-negative Matrix Factorization using gene expression data, International Conference on Information Communication and Embedded Systems (ICICES), 2013. 21-22. 2013, Chennai, India.

[9]   Swangnetr, M.,Kaber, D.B.,   Emotional State Classification in Patient–Robot Interaction Using Wavelet Analysis and Statistics-Based Feature Selection, IEEE Transactions on Human-Machine Systems (Volume:43,Issue:1),.2013.

[10] W Cai ;S. Chen ;D. Zhang. A Multi-objective Simultaneous Learning Framework for Clustering and Classification . IEEE Transactions on Neural Networks, Volume: 21 , Issue: 2 2010.

[11] E. R., Pfahringer, B., Holmes, G., Clustering for classification . Information Technology in Asia (CITA 11), 2011 7th International Conference on Digital Object Identifier: 10.1109/CITA.2011.5998839. 2011 , Page(s): 1 – 8.

[12] A. Shahi, R. Binti Atan and M-D. Nasir Sulaiman, An effective fuzzy c-mean and type-2 fuzzy logic for weather forecasting . Journal of Theoretical and Applied Information Technology. 2009, Vol. 5 Issue 5, p550 . Malaysia.

[13] Q. Liang and J. Mendel, Decision Feedback Equalizer for Nonlinear Time-Varying Channels Using Type-2 fizzy Adaptive Filters . Fuzzy Systems, 2000..

[14] G. Zhengetal, A Similarity Measure between Interval Type-2 Fuzzy Sets . Proceedings of the 2010 IEEE, International Conference on Mechatronics and Automation 2011.

[15] Kimito Funatsu and Kiyoshi Hasegawa, New fundamental technologies in data mining. First published January, 2011. Printed in India.

[16] Rui Xu, Donald Wunsch II, Survey of Clustering Algorithms . IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 16, NO. 3, 2005.

[17] Huaxiang Zhang , Jing Lu , Creating ensembles of classifiers via fuzzy clustering and deflection . Fuzzy Sets and Systems, Volume 161, Issue 13, 1 2010, Pages 1790–1802.

[18] Ondrej Linda, Milos Manic, General Type-2 Fuzzy C-Means Algorithm for Uncertain Fuzzy Clustering . Fuzzy Systems, IEEE Transactions. 13 2012, ISSN : 1063-6706.

[19] Frank, A. Asuncion, UCI Machine Learning Repository [http://archive.ics.uci.edu/ml], Irvine, CA: University of California, School of Informatics and Computer Science.

[20] R. Athauda, M. Tissera, C. Fernando. ” Data Mining Applications: Promise and Challenges”. Data Mining and Knowledge Discovery in Real Life Applications, ISBN 978-3-902613-53-0, pp. 438, 2009, I-Tech, Vienna, Austria.

[21] M. H. Fazel Zarandi, I. B. Turksen, O. Torabi Kasbi,” Type-2 fuzzy modeling for desulphurization of steel process”. Expert Systems with Applications 32 (2007) 157–171.

[22] D. Wu , J. Mendel.” Enhanced Karnik-Mendel Algorithms for Interval Type-2 Fuzzy Sets and Systems”. Fuzzy Information Processing Society, 2007. Annual Meeting of the North American.

[23] Der-Chen Lin, Miin-Shen Yang,” A similarity measure between type-2 fuzzy sets with its application to clustering”. Fourth International Conference on Fuzzy Systems and Knowledge Discovery, 2007

[24] Wen-liang Hung, Miin-shen Yang ,” Similarity Measures Between Type-2 Fuzzy Sets”. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. Vol. 12, No. 6 (2004) 827-841.

[25] Miin-Shen Yang, Der-Chen Lin. “On similarity and inclusion measures between type-2 fuzzy sets with an application to clustering”. Computers and Mathematics with Applications 57 (2009) 896_907.

[26] Hwang C.-M., Yang M.-S., Hung W.-L., “On similarity, inclusion measure and entropy between type-2 fuzzy sets”. International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems 2012. Volume 53, Issues 9–10.