

CLUSTER IDENTIFICATION FOR MICROARRAY GENE EXPRESSION DATA UNDER CONFLICT OF INTEREST

ANANDHAVALLI SUBRAMANIAN^{1,2} AND SRINIVASA KRISHNA SRIVATSA³

¹Department of Computer Science and Engineering
Jawaharlal Nehru Technological University
Kukatpally, Hyderabad 500085, India
{ anandhavalli0566; anandhi.bharat }@gmail.com

²Department of Master of Computer Applications
St. Joseph's College of Engineering
Chennai 600119, India

³Department of Computer Science and Engineering
Prathyusha Institute of Technology and Management
Poonamalee, Tiruvallur 602025, India
profsks@rediffmail.com

Received November 2013; revised November 2014

ABSTRACT. *In our former deal with conflict analysis, the conflicting information is evaluated by exploiting enhanced variant of FLEXFIS using overlapping clustering methodology. The execution of grouping methodology is assessed by employing two microarray gene expression datasets. Nevertheless, this method has achieved low grouping exactness in arranging the gene expression information. Subsequently, to enhance the gene quality order exactness and fortitude the conflicts, another conflict research method is proposed with modified fuzzy c-means clustering strategy. Our proposed strategy involves four phases viz., MFCM based grouping, fuzzy guidelines framing, conflict analysis and semi supervised learning. The MFCM clustering with our strategy ensures improvement in convergence speed with reduction in plasticity – stability dilemma. The fuzzy principles are created for the grouped information and accordingly the conflict methodology is performed to look at the information which is available in more than one cluster. The choice of MFCM strategy guarantees variation in convergence speed with diminishment in plasticity – constancy predicament. The usage result demonstrates the viability of proposed conflict analysis procedure in clustering the information in a characterized group. The execution of the proposed conflict procedure is assessed by leading different probes distinctive microarray gene expression datasets. In addition, the execution of the proposed strategy is contrasted with the current FLEXFIS and FLEXFIS with overlapping clustering methodologies. The analytic consequence demonstrates that our proposed strategy more precisely groups the gene data into their appropriate cluster or tenet than any of the other prevailing methodologies with high order precision.*

Keywords: Clustering, Fuzzy rules, Takagi Sugeno fuzzy model, Conflict analysis, Fuzzy c-means (FCM) – semi supervised learning

1. **Introduction.** Fuzzy logic or fuzzy set hypothesis was proposed by L. A. Zadeh in 1965. This has been effectively connected with numerous fields [1-5]. It is stably settled on numerical illustration of human speculation framework to manage instability [6] and is a quantitative methodology for managing the issues which possess fuzziness in intricate frameworks [3,7]. Fuzzy logic permits the objects in a set to be depicted as having a certain enrollment degree [8]. In fuzzy set hypothesis, a specific protest in a given set has a level of enrollment esteem in the limit between 0 (totally not in the set) and 1 (totally

in the set) [3] and is regularly alluded as multi-esteemed logic (0 to 1) in light of the fact that it encourages the meaning of intermediate values detached from the conventional assessments [9]. Simultaneously few applications of Fuzzy [25] frameworks are used for ecological control (air conditioning systems, humidifiers), residential apparatuses (washing machines, vacuum cleaners, toasters, microwave ovens, refrigerators), consumer electronics (TV, scanners, cameras, HI-FI frameworks) and auto frameworks (vehicle atmosphere control, programmed gearboxes, four-wheel steering, seat/mirror control frameworks) [10].

Fuzzy logic is utilized as a device for accomplishing intellectual control. The establishment of tenet based conduct is created by fuzzy logic. It is employed to encrypt the skill of a master into a standard base and recover the equivalent data in verdict determining [11]. Fuzzy system is one of the strategies to be utilized demonstrating nonlinear, indeterminate, and intricate frameworks [12] and those models have an outstanding framework portrayal ability than the linear systems [2]. An imperative normal for fuzzy models is partitioning the space of the framework variables into different fuzzy regions by employing fuzzy sets. In every locale, a tenet might be utilized to depict the qualities of the framework. A fuzzy system generally comprises a principle base with a guideline for every specific locale [12]. Fuzzy demonstrating is frequently employed within frameworks recognizable proof and control, flaw evaluation, characterization and decision supportive networks [5].

Takagi-Sugeno (TS) model [13] assumes a critical part in framework establishing and distinguishing proof among the various fuzzy displaying methods by consolidating the high estimation precision with phonetic interpretability [14,15]. This system is employed to disintegrate the data space of the nonlinear model into fuzzy subspaces and after that estimate the framework in every subspace by a basic linear regression model [16]. There are two phases included in the development of a TS model. During the initial phase, the fuzzy sets (enrollment capacities) in the guideline precursors are dictated by employing the skill of the strategy, or through some basic data driven strategies. In the next phase, the parameters of the resulting capacities are evaluated [15]. These strategies are habitually connected in a few modern applications and have the attired property to inexact any true happening nonlinear correlation to an assured level of perfection [17].

The conventional Fuzzy *c*-means (FCM) clustering strategy adopts Euclidean Distance to estimate the difference between objects. Appropriately the groups with convex forms possibly are found. In any case, it is hard to find the groups with spasmodic/irregular shapes; furthermore, the group is more delicate to the presence of noise and detached points [22-24].

The enduring assessment of the paper is composed as taking after. In Section 2, a concise audit is made about the current analysis task identified with Takagi Sugeno fuzzy model. Section 3 depicts the proposed conflict strategy with modified fuzzy *c*-means clustering strategy with essential numerical detailing of the system concerning the micro cluster gene representation data sets. In this way, Section 4 examines about the execution and similar results while Section 5 deliberates about the inference and forthcoming effort.

2. Literature Survey. Several analyses are accessible for Takagi-Sugeno fuzzy frameworks and fuzzy clustering strategies. Few of them are investigated in what takes after.

In 2007, Y. Chen et al. [18] have proposed a programmed method for developing progressive Takagi-Sugeno fuzzy frameworks (TS-FS). The progressive structure was developed by employing a probabilistic incremental program evolution (PIPE) with particular guidelines. The calibrating of the If-then guideline's constraints was programmed in the structure and it was fulfilled by evolutionary programming (EP). Their system interleaved both PIPE and EP advancements. Their technique began with arbitrary structures and

standards/constraints. It attempted to enhance the progressive structure by the tweaks and tenets parameters. This loop proceeded awaiting a discerning result (progressive TS-FS model) was discovered or a period utmost was arrived at. Their proposed progressive TS-FS was assessed by employing definite benchmark applications named recognizable nonlinear frameworks, forecast of the Mackey-Glass chaotic time-series and some order issues. The analysis has indicated high precision execution of their progressive TS-FS system when contrasted with other neural systems and fuzzy frameworks.

In 2008, E. D. Lughofer [17] have suggested about a strategy for incremental learning of a particular manifestation of Takagi-Sugeno fuzzy frameworks proposed by Wang and Mendel in 1992 [30]. This methodology did not just incorporate the adjustment of linear parameters showing up in the guideline consequents, additionally the incremental learning of principle parameters showing up in the participation capacities (fuzzy sets). An altered variant of vector quantization was misused for standard advancement and an incremental learning of the tenets purpose segments. These adjustments have been utilized to create groups focused around the nature, dissemination, and eminence of fresh information and an option technique for selecting the winning cluster (guideline) in every incremental learning phase. Precursor and ensuing learning were linked in a constant way to acquire the ideal parameter set at minimum squares routine. This methodology was assessed and contrasted with regular cluster strategies focused around high-dimensional information records at motor test benches and at rolling plants. Also, this technique was contrasted with an explanatory corporal strategy.

In 2009, C.-J. Lin et al. [19] have anticipated a functional-link established neural fuzzy system (FLNFN) model. This model was focused around a hybrid of cooperative particle swarm optimization and cultural strategy and hence it was called cultural cooperative particle swarm optimization (CCPSO). Their planned CCPSO system has utilized cooperative conduct among various swarms and expanded the worldwide inquiry limit utilizing the conviction space. The conviction space was the data store in which the people could store their encounters so that different people could gain from them in an ultimate way. This strategy employs the functional link neural systems as the ensuing piece of the fuzzy principles and it was utilized to produce the resulting piece of a nonlinear unification of input variables. At last, their FLNFN strategy with CCPSO (FLNFN-CCPSO) was established in a few prescient applications. The analysis has exposed that their proposed CCPSO system performed well in anticipating the time arrangement issues.

In 2010, C.-F. Juang et al. [20] have proposed a hierarchical cluster-based multi-species particle-swarm optimization (HCMSPSO) strategy for fuzzy framework optimization. This methodology is employed to acquire Takagi-Sugeno-Kang (TSK) sort fuzzy standards with high correctness. In the HCMSPSO-outlined fuzzy system (FS), each one tenet was characterized with its own particular fuzzy sets, which inferred that the quantity of fuzzy sets for each one input variable is equivalent to the quantity of fuzzy standards. A swarm in HCMSPSO was clustered into different species at an upper progressive range, and every species was further grouped into various subspecies at a lower various leveled range. For an FS comprised of r standards, r species (swarms) were structured in the upper level and every species might be improved by a solitary fuzzy standard. At first, there were no species in HCMSPSO. An online cluster based strategy was employed to produce new species (fuzzy guidelines) consequently. In the lower layer, subspecies inside the identical species were structured adaptively in every emphasis amid the particle up gradation. A few simulations were led to confirm HCMSPSO execution and their system was contrasted with further neural learning, genetic, and PSO strategies to exhibit the prevalence of HCMSPSO accomplishment.

In 2011, O. Cordon [21] has presented the requirement for exchanging off interpretability and precision in fuzzy frameworks. The requirement for correct and human-intelligible fuzzy frameworks assumed a key part in Zadeh and Mamdani's fundamental thoughts and framework ID strategies. Prior to the appearance of delicate processing, precision continuously turned into the principle concern of fuzzy model. The developers were to make the ensuing fuzzy frameworks get quicker to black-box models, for example, neural systems. Specifically, the utilization of hereditary fuzzy frameworks has been generally stretched out to their innate adaptability and their capacity to mutually consider various advancement principles. This effort established a survey on the most illustrative genetic fuzzy frameworks depending on Mamdani-sort fuzzy principle based frameworks to get interpretable semantic fuzzy models with a better precision.

Motivation of the work

Microarray gene expression data are vast that take part in the majority of the basic practices. DNA micro arrays are utilized to check the improvement/advancement of gene interpretations. Thus, proper clustering of such outflow information is obliged to recognize and perform these exercises. The principle destination of grouping is to segment a whole array of gene expression into a few segments or groups that have comparable conduct. Grouping system might be utilized to break down the characterization of such capacious information into the proper clusters. A few techniques, for example, hierarchical grouping, partition clustering is accessible to control such information. Here we have employed a strategy to abide the approaching microarray gene expression sets into the proper single cluster. Be that as it may, the impact of a specific gene that helps the whole state of cluster must additionally be considered. Henceforth, we employ a fuzzy grouping system, all the more particularly, modified fuzzy c -means strategy. The fuzzy c -means strategy focused around Lagrangian multiplier permits us to determine the neighborhood ideal result that serves to settle on different start procedures of approaching data focuses as micro array gene expression sets. When we determine entomb affiliations/inter associations amongst different gene expression that structure some piece of one/more clusters, we need to do the conflict analysis and choose which group is the best suited one for the approaching information. In our prior work that managed conflict strategy, the conflict information in micro cluster gene expression is evaluated by misusing enhanced form of FLEXFIS with overlapping clustering methodology. This system performs in two phases to be more specific, FLEXFIS with overlapping clustering approach and conflict analysis. Initially the input information is grouped into their separate groups by the FLEXFIS with overlapping clustering methodology. Additionally, this method has achieved 72% general accuracy in grouping the two microarray gene expression data. Be that as it may, this arrangement precision is deficient. Henceforth, another conflict analysis system is proposed to enhance the grouping accuracy in this proposed paper.

3. Proposed Modified FCM Based Conflict Analysis Technique. In our proposed strategy, the conflict methodology is done with the assistance of modified FCM and Takagi-Sugeno fuzzy framework. At first the information is clustered by employing the modified FCM system focused around Mahalanobis distance algorithm that is employed to coordinate the groups. Such clustering of information is obliged on the grounds that these information may be associated and included in performing comparative natural capacities. It is very normal that the protein products of qualities are included in numerous biological methodologies. Accordingly, such qualities may be corresponded in performing distinctive capacities under various conditions or circumstances. Such a conduct cannot be settled by employing conventional clustering methodologies since they are not adaptable enough to characterize, control and evaluate the conduct. Thus, modified

fuzzy c -means is ended up being better over other clustering systems as far as partitioning proficiency, convergence speed and diminishment in plasticity – dependability issue. An important, proper peculiarity vector space is utilized for the parceling method. The execution of the modified fuzzy c -means relies on the choice of beginning group focuses or participation values. In the event that a better starting focus that is near the real latter cluster could be found, the execution of the strategy will be exceptionally proficient. It will likewise bring about high level of linking to the result space. The fuzzy standards are produced for the clustered data qualities to recognize the proper expediency of the created micro array gene expression data sets. The proposed framework chiefly involves four phases such as (i) clustering by MFCM (ii) fuzzy principles generation (iii) conflict analysis and (iv) regulated learning. These four phases are continuously performed and the more faultless results are acquired. Such results are scrutinized in Sections 3.1, 3.2 and 3.3 individually. The essential structure of our proposed strategy with MFCM based conflict strategy is indicated in Figure 1. Despite the fact that our strategy makes operation of modified fuzzy c -means strategy, it likewise considers semi-administered adapting in recognizing the proper group for the approaching input data.

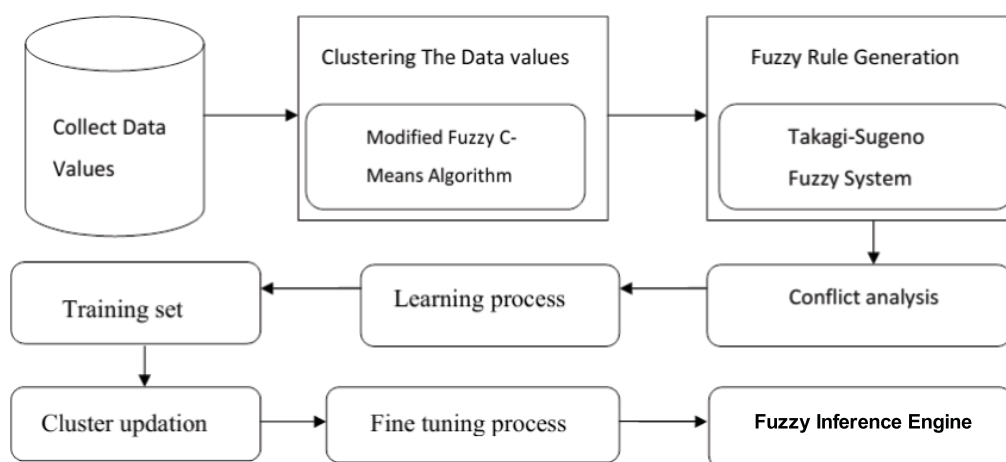


FIGURE 1. Structure of our proposed Resolve_Conflict MFCM based conflict analysis technique

The pseudo code of our Resolve_Conflict algorithm with modified fuzzy c -means algorithm is stated as in what follows:

Resolve_Conflict Algorithm

Input: Micro array gene Data Set

Output: Selection of appropriate clusters

STEPS

1. Extract the dissemination of micro array gene data sets in the arrangement of several parameters in the feature space.
2. Partition the input space into smaller local regions, recognized as clusters.
3. Analyze the initialization method by embracing an arbitrary transfer of membership functions.
4. Allocate the membership values appropriately.
5. Launch the data set framework and fuzzifier relationship.
6. Assess the ideal value for fuzzifier.
7. Determine the genes in a proper manner that they are compactly related to the appropriate clusters.

8. Portray the confined possessions by normalizing the membership functions in the fuzzy set.
 9. Employ sigmoidal threshold to describe the worldwide possessions of the learning process.
 10. Apprise the training set as and when fresh data points are determined.
 11. Employ reworking of vigilance constraint for cluster updates.
 12. Fine tune the vigilance constraint by setting the threshold value.
-

3.1. Clustering by MFCM. FCM is equivalent as the k -means classifier which permits one bit of information vital fit in with more than one cluster. A class membership value is pointed out to every dynamic data, contingent upon the similarity of the information crucial to a specific class association to different classes. In our proposed philosophy, every variable with each group utilizes an enrollment work that communicates the variable's terminating quality of the affiliation. This produces sets of non-selective clusters that permit qualities to have membership works in various groups as opposed to in precisely stand out cluster. The introduction methodology of the strategy depends on irregular task of enrollment of genes to the clusters. In this way, the produced genes are conflicting in separating the groups. An alternate condition is that the quantity of anticipated that groups which will be framed must be from the earlier known. Thus the requirement for investigating the conflicts employing modified fuzzy c -means strategy with semi managed learning methodology. When the groups are found, biological learning might be utilized to scan for proof of the procedure-based relationship inside the clusters. However, genes of abnormal capacities are allotted to groups with genes of known capacities, and conflicts emerges in such a path, to the point that genes with known capacities are assigned to different clusters. This even restricts the capacity to discover new capacities of legitimately expressed genes by connecting with other known capacities. Subsequently, to enhance the grouping exactness, a modified FCM strategy with new target capacity is abused in our paper. In this method, the destination capacity " O " is characterized as,

$$O = \sum_{i=1}^N \sum_{j=1}^{ce} m_{ij}^{\phi} \|x_i - Ce_j\|^2 \quad (1)$$

This standard FCM objective function in Equation (1) is modified by including the constant threshold value. Thus the modified objective function in MFCM is in what follows.

3.1.1. Modified fuzzy c -means clustering (MFCM). The input data values are clustered by employing the modified FCM strategy. The input data values are clustered in one cluster which provides identical values. The structures clustering strategy by MFCM technique objective function is indicated by,

$$O = \sum_{k=1}^K \sum_{r=1}^R \left[(1 - \beta) m_{kr}^{\phi} \|i_k - c_r\|^2 \right] \quad (2)$$

The MFCM objective capacity is adjusted by reproducing the steady esteem β value with the membership capacity. In Equation (2), m_{kr} is the level of membership i_k in the cluster r , i_k is the data and c_r is the centroid value of the cluster r , and $\|*\|$ is the similitude between any measured information and the group focus. The information is apportioned by the fuzzy capacity and the membership and cluster focus are upgraded

by,

$$m_{kr} = \frac{1}{\sum_{s=1}^R \left(\frac{\|i_k - c_r\|}{\|i_k - c_s\|} \right)^{\frac{2}{m-1}}} \tag{3}$$

$$c_r = \frac{\sum_{k=1}^K m_{kr}^\phi \cdot i_k}{\sum_{k=1}^K m_{kr}^\phi} \tag{4}$$

The method is repetitive when the iteration is grasped. The centroid values from the MFCM clustering are indicated as $C = \{c_1, c_2, \dots, c_R\}$, where R is the number of clusters.

3.2. Fuzzy rules generation. The process of fuzzy rules generation by the Takagi-Sugeno fuzzy system is defined with number of input variables (n_1, \dots, n_p) , a single output variable o . The rule C can be defined as stated below

$$f(n) = o = \sum_{i=1}^C l_i \psi_i(n) \tag{5}$$

$$\psi_i(n) = \frac{\exp \left[-\frac{1}{2} \sum_{x=1}^p (n_x - c_{ix})^2 / \sigma_{ix}^2 \right]}{\sum_{k=1}^C \exp \left[-\frac{1}{2} \sum_{x=1}^p (n_x - c_{kx})^2 / \sigma_{kx}^2 \right]} \tag{6}$$

$$l_i = w_{i0} + w_{i1}n_1 + w_{i2}n_2 + \dots + w_{ip}n_p \tag{7}$$

The fuzzy membership and ensuing capacities are delineated in Equations (6) and (7). In Equation (6), symbol n_x indicates the x^{th} input variable, c_{ix} is the center, and σ_{ix} is the width of the Gaussian fuzzy set in the x^{th} premise part of the i^{th} rule. At the point when examining this plan of a Takagi-Sugeno fuzzy framework, we understand that chiefly, three separate portions may be developed, that is, adapted in an incremental way: direct ensuing parameters showing up in the standards consequents as output weights $(w_{i0}, w_{i1}, \dots, w_{ip})$, nonlinear principle constraints showing up in the input membership based on the centers (c_{ix}) and widths (σ_{ix}) and the principle base, concerning the quantity of principles C and the quantity of fuzzy sets for every input measurement.

3.3. Conflict analysis. We dissect the data focuses which are included in the conflict process, i.e., the data focuses that have identical property are assembled into one or more groups. All the more particularly, the conflict may happen by employing the microarray gene expression datasets. Subsequently the genes have comparable properties, and there is a more noteworthy opportunity to the genes being grouped in one or more clusters. The methodology of over grouping does not make the exact clustering results. In the prevailing FLEXFIS and FLEXFIS with overlapping clustering methodologies, execution lacks in the conflict analysis and grouping methodology. In our proposed strategy with MFCM based conflict strategy the information is clustered in a more effective manner. In conflict analysis, we test if any data focuses or gene qualities are available in one or more clusters or standard. Our proposed MFCM has given low conflict values when clustering the data qualities contrasted with other prevailing methodologies. The calibrating methodology is carried out by the utilization of vigilance parameter that chooses the limit worth to be 0.05 as the membership cutoff quality is naturally sensible, as it is higher than the consistently dispersed membership estimation of (0.03125). The decision of this membership cutoff quality is additionally defended by the perception that around 96% of the genes had membership estimations of ≤ 0.05 , while around 3.6% of the qualities had membership estimations of ≥ 0.1 . Subsequently this strategy is by outline determinedly predisposition towards grouping focused around the earlier skill which is employed, and it is most likely

Fuzzy rules	
1	$y = \exp\left(-\frac{1}{2}\right) * [((x_1 - 0.093767)^2 / (0.1)^2) + ((x_2 - 0.0861)^2 / (0.1)^2) + ((x_3 - 0.04)^2 / (0.1)^2) +$ $((x_4 - 0.12515)^2 / (0.1)^2) + ((x_5 - 0.20235)^2 / (0.1)^2) + ((x_6 - 0.090868)^2 / (0.1)^2) +$ $((x_7 - 0.052989)^2 / (0.1)^2) + ((x_8 - 0.14737)^2 / (0.1)^2) + ((x_9 - 0.14993)^2 / (0.1)^2)] +$ $\sum_{j=0}^{37} \frac{(x_{10+j} * 10^{-7} - c_0)^2}{(d_0 + j * 10^{-2})^2}$
2	$y = \exp\left(-\frac{1}{2}\right) * [((x_1 - 0.048)^2 / (0.18)^2) + ((x_2 - 0.049)^2 / (0.19)^2) + \sum_{j=0}^{12} \frac{(x_{3+j} * 10^{-7} - c_1)^2}{(d_1 + j * 10^{-2})^2} +$ $\sum_{j=0}^{22} \frac{(x_{13+j} * 10^{-7} - c_2)^2}{(d_2 + j * 10^{-2})^2} + \sum_{j=0}^{32} \frac{(x_{23+j} * 10^{-7} - c_3)^2}{(d_3 + j * 10^{-2})^2} + \sum_{j=0}^{37} \frac{(x_{33+j} * 10^{-7} - c_4)^2}{(d_4 + j * 10^{-2})^2}]$
3	$y = \exp\left(-\frac{1}{2}\right) * [((x_1 - 0.11717)^2 / (0.1)^2) + ((x_2 - 0.0048137)^2 / (0.1)^2) +$ $((x_3 - 0.023301)^2 / (0.1)^2) + ((x_4 - 0.10348)^2 / (0.1)^2) + ((x_5 - 0.093427)^2 / (0.1)^2) +$ $((x_6 - 0.14017)^2 / (0.1)^2) + ((x_7 - 0.10452)^2 / (0.10834)^2) + ((x_8 - 0.3771)^2 / (0.1)^2) +$ $((x_9 - 0.144)^2 / (0.1)^2)] + \sum_{j=0}^{37} \frac{(x_{10+j} * 10^{-7} - c_5)^2}{(d_5 + j * 10^{-2})^2}$

FIGURE 2. Rules generated for ALL&AML and AD&NL datasets

not best suited for clustering where affiliations focused around that specific sort of former information is not straightforwardly identified with the evaluation method.

3.4. Supervised learning. The unreliable genes that fit in with the non-selective groups are given as input to the learning procedure. These genes are contrasted and the preparation set so the winning cluster is recognized by embracing closest neighborhood procedure. Along these lines, even the conflicting genes are additionally placed in the proper, significant clusters to enhance the precision of our strategy. The principles which choose the proper identification of the group are again put away in the fuzzy inference motor with the goal that the advance generation of the standards is additionally enhanced significantly. The execution of the vigilance parameter is moreover measured to calibrate the whole methodology of cluster recognition.

4. Experimental Results and Discussion. The offered MFCM based conflict analysis system is actualized in the operating stage of MATLAB (variant 7.12). The proposed conflict strategy execution is assessed with two various microarray gene datasets, namely, ALL&AML [22] and AD&NL [23]. The AD&NL dataset comprises 675 genes and 156 samples. Between these 156 samples, 1 to 139 samples are appropriate to the AD class type and 140 to 156 samples are in the right place to the NL class type. The ALL&AML dataset is of dimension 7192×38. In this data set, 1 to 26 gene column data are ALL and 26 to 38 gene column data are AML.

At first, the dataset gene qualities are clustered by applying MFCM and that grouped genes relating fuzzy tenets are created by the Takagi-Sugeno fuzzy framework. The grouping process on these gene interpretation datasets produces the conflicts between the clustered gene information. So there is a need to dissect the conflict between the clustered gene

qualities. We lead the conflict analysis handle on both ALL&AML and AD&NL datasets fuzzy guidelines grouping results. The proposed method likewise utilizes semi regulated learning process that is utilized for taking a choice as a part of recognizing the partaking group. Such a choice is brought with the assistance of the preparation set of effectively prevailing information. If there should arise an occurrence of new group being produced, such cluster recognition and redesigning of the aggregate number of groups is carried out by the strategy. The proposed MFCM with resolve_conflict procedure execution is investigated in clustering the gene qualities and is enhanced by the adjusting process that is assessed by the vigilance parameter. The anticipated conflict analysis system’s execution is investigated by directing four separate probes both gene outflow datasets and their grouped results from each one analysis are acquired. Additionally, our proposed MFCM based conflict analysis procedure execution is contrasted with the prevailing FLEXFIS and FLEXFIS with overlapping clustering methodologies. The clustering consequences from our proposed and prevailing procedures are demonstrated in Table 1.

Our suggested procedure researches the execution by employing the numerical measures which is given in [24] and a similar analysis is additionally completed with the current strategies. The consequences of such conflict analysis are expressed in Table 3. In the execution analysis, the ALL information is partitioned into four and AML information into four; absolutely we have 8 clusters of ALL&AML data and the AD&NL dataset likewise partitioned into four clusters. The numerical measures and the consequences of such

TABLE 1. Clustering results of proposed and existing FLEXFIS and FLEXFIS with overlapping clustering approach technique: (a) ALL&AML, (b) AD&NL

Experiments	Reliable Cluster Results	Obtained Result		
		Proposed MFCM Based Conflict Analysis Technique	FLEXFIS with Overlapping Clustering Approach	Existing FLEXFIS Approach
1	1-6	1,2,4,5,8	1,4,5,9	1,5
	7-9	3,6,7,9	2,3,6,7,8	2
2	1-6	1,2,3,4,5,8	1,2,3,4,5,6,7,8,9	1
	7-9	6,7,9	0	2
3	1-6	1,4,5	1,2,3,4,5,6,7,8,9	1,5,6,7,8,9
	7-9	2,3,6,7,8,9	0	2
4	1-6	1,2	1,4,7,9	1,7,9
	7-9	3,4,5,6,7,8,9	2,3,5,6,8	2

(a)

Experiments	Reliable Cluster Results	Obtained Result		
		Proposed MFCM Based Conflict Analysis Technique	FLEXFIS with Overlapping Clustering Approach	Existing FLEXFIS Approach
1	1-34	1-33,36,37,38	1-38	1
	35-38	35,34	0	2,3
2	1-34	1-38	1-38	1,9,23,24,25,32
	35-38	0	0	2
3	1-34	1-38	1-38	1,7:38
	35-38	0	0	2
4	1-34	1-38	1-38	1
	35-38	0	0	2

(b)

TABLE 2. Proposed conflict analysis technique TP, TN, FP and FN values from four different experiments: (a) ALL&AML, (b) AD&NL datasets

Experiments	TP	FP	TN	FN
1	4	1	2	2
2	5	1	2	1
3	3	0	3	3
4	2	0	3	4

(a)

Experiments	TP	FP	TN	FN
1	33	3	1	1
2	34	4	0	0
3	34	4	0	0
4	34	4	0	0

(b)

TABLE 3. Proposed MFCM based conflict analysis technique for different experiments from ALL&AML

Experiments	Sensitivity	FPR	Accuracy	PPR	NPV	FPR	MCC
1	67	33	67	80	50	20	32
2	83	33	78	83	67	17	50
3	50	0	67	100	50	0	50
4	33	0	56	100	43	0	38
Average	58.25	16.5	67	90.75	52.5	9.25	42.5

TABLE 4. Proposed MFCM based conflict analysis technique for different experiments from AD&NL datasets

Experiments	Sensitivity	FPR	Accuracy	PPR	NPV	FPR	MCC
1	97	75	89	92	50	83	30
2	100	100	89	89	0	11	0
3	100	100	89	89	0	11	0
4	100	100	89	89	0	11	0
Average	99.25	93.75	89	89.75	12.5	29	7.5

information sets from our proposed and prevailing systems are demonstrated in Tables 4 and 5. These numerical measures are ascertained by employing the four parameters like TP, TN, FP and FN. At first these parameters qualities ascertained from the clustering results are given in Table 2.

Furthermore, the performance evaluation graphs of our planned and prevailing strategies for the ALL&AML and AD&NL datasets are exposed in Figure 3.

The performance assessment chart of Figure 3, demonstrates that our proposed MFCM based conflict analysis system has achieved exclusive performance and sensitivity values. In ALL&AML dataset the planned strategy precision is higher than the modern technique yet in sensitivity degree the proposed system has accomplished lower sensitivity than the current FLEXFIS with overlapping clustering methodology. In AD&NL dataset, the proposed system has gained higher sensitivity and precision esteem than the prevailing FLEXFIS with overlapping clustering methodology. The low execution of sensitivity

TABLE 5. Existing conflict analysis technique with FLEXFIS-overlapping clustering approach for 4 different experiments from (a) ALL&AML (b) AD&NL datasets

Experiments	Sensitivity	FPR	Accuracy	PPR	NPV	FPR	MCC
1	50	33.3	56	75	40	25	14.1
2	100	100	67	67	0	33	0
3	100	100	67	67	0	33	0
4	0	66.7	33	0	20	0	0
Average	63	75	56	52	15	23	4

(a)

Experiments	Sensitivity	FPR	Accuracy	PPR	NPV	FPR	MCC
1	100	100	89	89	0	11	0
2	0	100	89	0	0	0	0
3	100	100	89	89	0	11	0
4	100	100	89	89	0	11	0
Average	75	100	89	67	0	8	0

(b)

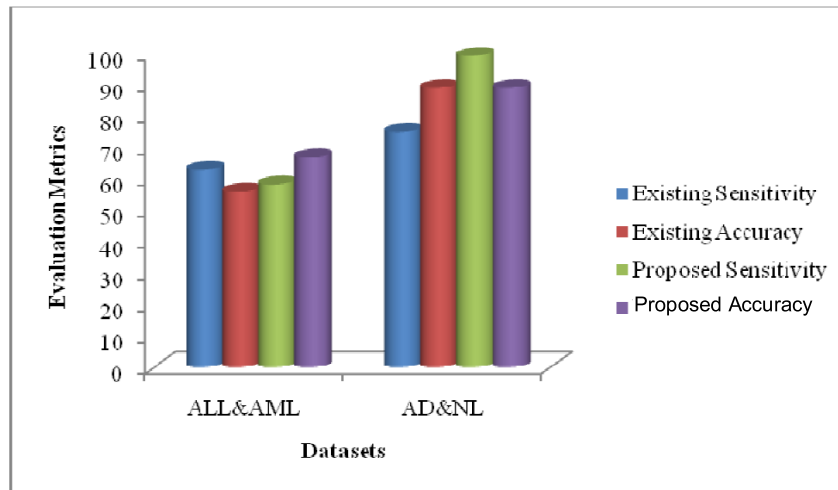


FIGURE 3. Performance evaluation graph

degree in ALL&AML dataset does not influence the proposed method execution on the grounds that the anticipated strategy sensitivity degree execution in AD&NL dataset is much strayed from that low accomplishment.

By comparison, the accomplishment of our proposed methodology is assessed by employing the prevailing strategy characterization accuracy. The classification correctness execution estimations of our proposed, prevailing FLEXFIS and FLEXFIS with overlapping clustering strategies consequences for the dataset ALL&AML and AD&NL are specified in Tables 6 and 7. Besides, the assessment diagram of proposed, prevailing FLEXFIS and FLEXFIS with overlapping clustering systems precision accomplishment is delineated in Figure 4.

As might be perceived from Figure 4, the proposed methodology has given higher grouping correctness than the prevailing strategies. Our proposed system has accomplished 67% of general arrangement precision execution in organizing the ALL&AML datasets genes qualities, yet the current FLEXFIS and FLEXFIS with overlapping clustering methods

TABLE 6. ALL&AML dataset classification accuracy of proposed, existing FLEXFIS and FLEXFIS with overlapping clustering techniques

Experiments	Proposed MFCM Based Conflict Analysis Technique	FLEXFIS with Overlapping Clustering Approach	Existing FLEXFIS Approach
1	67	56	22
2	78	67	11
3	67	67	22
4	56	33	11

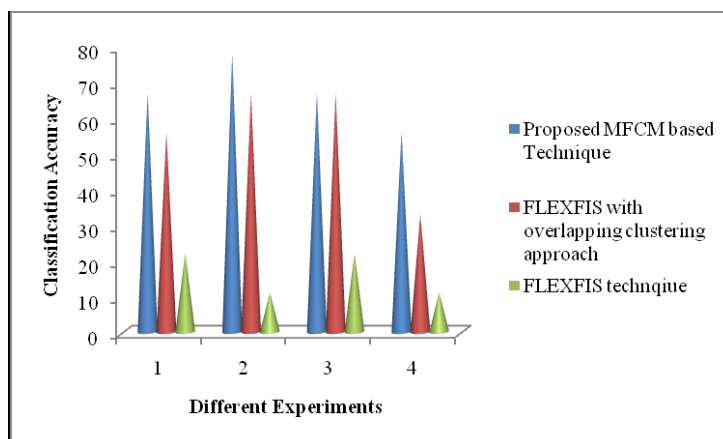
TABLE 7. AD&NL dataset classification accuracy of proposed, existing FLEXFIS and FLEXFIS with overlapping clustering techniques

Experiments	Proposed MFCM Based Conflict Analysis Technique	FLEXFIS with Overlapping Clustering Approach	Existing FLEXFIS Approach
1	89	89	3
2	89	89	16
3	89	89	76
4	89	89	3

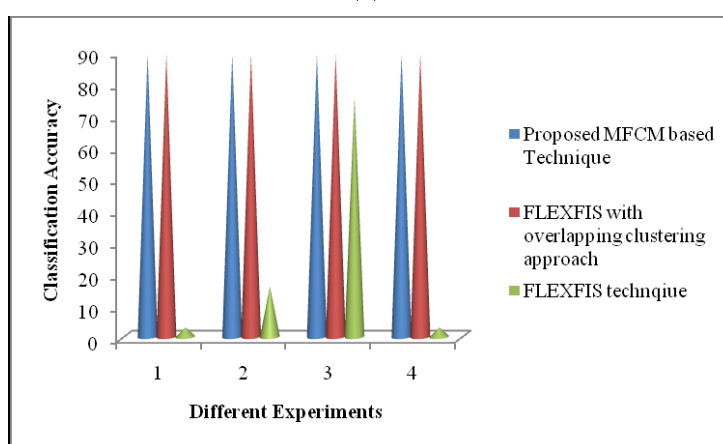
have achieved the characterization exactness of 56% and 17% individually. In Figure 4(b), in AD&NL dataset planning our proposed and prevailing FLEXFIS with overlapping clustering strategy has accomplished identical characterization perfection as 89% and those other performances are higher than the FLEXFIS method. Thus our proposed methodology has given higher characterization in both ALL&AML and AD&NL datasets than the prevailing strategy. It creates all the more organically significant consequences. It takes out the extra manual exertion to recognize capacities connected with the clusters. We utilize semi-managed to figure out how to enhance the execution of grouping and clustering.

5. Conclusion and Future Work. The paper has depicted the extent of investigating conflicts between the gene information by employing proposed MFCM based conflict analysis method. In the proposed method, MFCM is employed to cluster the approaching gene data with particular to their clusters and the fuzzy standards are produced for the clustered gene data values. To dissect the clustering execution a conflict analysis procedure is achieved above the clustered gene data. The proposed conflict analysis system execution is assessed by the two microarray gene expression datasets to be specific ALL&AML and AD&NL. The consequence of our proposed method from this both datasets demonstrates that the proposed MFCM based conflict analysis system gets higher quality order precision than the prevailing FLEXFIS and FLEXFIS with overlapping clustering methodology. Henceforth, it is demonstrated that our planned conflict analysis strategy with MFCM clustering strategy all the more definitely groups the information into their ensuing clusters with high quality order precision. It gives naturally serious characterization of appropriating clusters to the gene data sets. The proposed system has effective applications in a few fields, fiscal and restorative among others. It is so impulsive feasible to conclude that its application to the conflict analysis can fetch optimistic and expedient consequences.

The forthcoming effort may ensue in the subsequent orders.



(a)



(b)

FIGURE 4. Performance comparison graph: (a) ALL&AML dataset, (b) AD&NL dataset

- i) Unsupervised learning algorithm can be established for assigning the weight on identical micro array gene data sets to minimize the conflicts.
- ii) Enhancing strategies can be implemented without aggregating the intricacy of cluster recognition.
- iii) Estimation of gene functions can be established.
- iv) The uncertainty in clustering can be condensed to a definite range.
- v) The quantity of clusters can be spontaneously assessed by means of FLAME to find non-linear interactions and non-globular clusters.

REFERENCES

- [1] L. A. Zadeh, Fuzzy sets, *Information and Control*, 1965.
- [2] E. Kim, M. Park, S. Kim and M. Park, A transformed input-domain approach to fuzzy modeling, *IEEE Trans. Fuzzy Systems*, vol.6, no.4, pp.596-604, 1998.
- [3] V. Kumar and R. R. Joshi, Hybrid controller based intelligent speed control of induction motor, *Journal of Theoretical and Applied Information Technology*, vol.3, no.1, pp.71-75, 2005.
- [4] A. Mustafa, A.-H. A. Hashim, O. Khalifa and S. A. Hamed, Adaptive emotional personality model based on fuzzy logic interpretation of five factor theory, *International Journal of Signal Processing*, vol.2, no.4, pp.1-9, 2008.
- [5] H. Roubos and M. Setnes, Compact and transparent fuzzy models and classifiers through iterative complexity reduction, *IEEE Trans. Fuzzy Systems*, vol.9, no.4, pp.516-524, 2001.

- [6] A. H. Isik and I. Guler, Novel approach for optimization of cell planning and capacity, *International Journal of Communications*, vol.3, no.1, pp.1-8, 2009.
- [7] C. E. Woodcock and S. Gopal, Fuzzy set theory and thematic maps: Accuracy assessment and area estimation, *Int. J. Geographical Information Science*, vol.14, no.2, pp.153-172, 2000.
- [8] M. Zeraoulia, A. Mamoune, H. Mangel and M. E. H. Benbouzid, A simple fuzzy logic approach for induction motors stator condition monitoring, *J. Electrical Systems*, vol.1, no.1, pp.15-25, 2005.
- [9] K. Angayarkkani and N. Radhakrishnan, Efficient forest fire detection system: A spatial data mining and image processing based approach, *IJCSNS International Journal of Computer Science and Network Security*, vol.9, no.3, pp.100-107, 2009.
- [10] D. Sharma, Designing and modeling fuzzy control systems, *International Journal of Computer Applications*, vol.16, no.1, pp.46-53, 2011.
- [11] A. Rajagopalan, G. Washington, G. Rizzoni and Y. Guezennec, *Development of Fuzzy Logic and Neural Network Control and Advanced Emissions Modeling for Parallel Hybrid Vehicles*, Subcontract Report, Center for Automotive Research, 2003.
- [12] M. Setnes, R. Babuska, U. Kaymak and H. R. van Nauta Lemke, Similarity measures in fuzzy rule base simplification, *IEEE Trans. Systems, Man, and Cybernetics – Part B: Cybernetics*, vol.28, no.3, pp.376-386, 1998.
- [13] T. Takagi and M. Sugeno, Fuzzy identification of systems and its application to modeling and control, *IEEE Trans. Systems, Man and Cybernetics*, vol.15, no.1, pp.116-132, 1985.
- [14] E. Lughofer, E. Hullermeier and E.-P. Klement, Improving the interpretability of data-driven evolving fuzzy systems, *Proc. of Joint the 4th Conference of the European Society for Fuzzy Logic and Technology and the 11th Rencontres Francophones sur la Logique Floue et ses Applications EUSFLAT*, pp.28-33, 2005.
- [15] J. Abonyi, R. Babuska and F. Szeifert, Modified Gath-Geva fuzzy clustering for identification of Takagi-Sugeno fuzzy models, *IEEE Trans. Systems, Man, and Cybernetics – Part B: Cybernetics*, vol.32, no.5, pp.612-621, 2002.
- [16] J. Abonyi, J. A. Roubos, M. Oosterom and F. Szeifert, Compact TS-fuzzy models through clustering and OLS plus FIS model reduction, *Proc. of the 10th IEEE International Conference on Fuzzy Systems*, Melbourne, Australia, pp.1420-1423, 2001.
- [17] E. D. Lughofer, FLEXFIS: A robust incremental learning approach for evolving Takagi-Sugeno fuzzy models, *IEEE Trans. Fuzzy Systems*, vol.16, no.6, pp.1393-1410, 2008.
- [18] Y. Chen, B. Yang, A. Abraham and L. Peng, Automatic design of hierarchical Takagi-Sugeno type fuzzy systems using evolutionary algorithms, *IEEE Trans. Fuzzy Systems*, vol.15, no.3, pp.385-397, 2007.
- [19] C.-J. Lin, C.-H. Chen and C.-T. Lin, A hybrid of cooperative particle swarm optimization and cultural algorithm for neural fuzzy networks and its prediction applications, *IEEE Trans. Systems, Man, and Cybernetics – Part C: Applications and Reviews*, vol.39, no.1, pp.55-68, 2009.
- [20] C.-F. Juang, C.-M. Hsiao and C.-H. Hsu, Hierarchical cluster-based multispecies particle-swarm optimization for fuzzy-system optimization, *IEEE Trans. Fuzzy Systems*, vol.8, no.1, pp.14-26, 2010.
- [21] O. Cordon, A historical review of evolutionary learning methods for Mamdani-type fuzzy rule-based systems: Designing interpretable genetic fuzzy systems, *International Journal of Approximate Reasoning*, vol.52, no.6, pp.894-913, 2011.
- [22] S. L. Teng, J. Y. Yang and L. J. Wang, Genome-wide prediction and analysis of human tissue-selective genes using microarray expression data, *BMC Medical Genomics*, vol.6, 2013.
- [23] L. Wang, K. Srivastava and C. E. Schwartz, Microarray data integration for genome-wide analysis of human tissue-selective gene expression, *BMC Genomics*, vol.11, 2010.
- [24] D. Dembele and P. Kastner, Fuzzy *c*-means method for clustering microarray data, *Bioinformatics*, vol.19, no.8, pp.973-980, 2003.
- [25] F. Jiang, F. Chen and C. Wen, Fuzzy control for road intersection signal with bus priority, *Aloy Journal of Soft Computing and Applications*, vol.1, no.1, 2014.
- [26] S. Anandhavalli and S. K. Srivatsa, Improved Takagi-Sugeno fuzzy modeling with FLEXFIS-overlapping clustering approach for efficient classification under conflicts of interest, *International Review on Modeling and Simulation*, vol.6, no.2, pp.654-662, 2013.
- [27] *ALL/AML Datasets*, <http://www.broadinstitute.org/cancer/software/genepattern/datasets/>.
- [28] *AD/NL Datasets*, <http://www.broadinstitute.org/cgi-bin/cancer/datasets.cgi>.
- [29] *Statistical Measures*, http://en.wikipedia.org/wiki/Sensitivity_and_specificity.
- [30] L.-X. Wang and J. M. Mendel, Generating fuzzy rules by learning from examples, *IEEE Trans. Systems, Man and Cybernetics*, vol.22, no.6, pp.1414-1427, 1992.