FUZZY CLUSTERING

Boldeanu Silviu
Faculty of Computer Science
Babes-Bolyai University
Email: silviu_boldeanu@yahoo.com

Abstract

In data mining, the conventional clustering algorithms have difficulties in handling the challenges posed by the collection of natural data which is often vague and uncertain. Fuzzy clustering methods have the potential to manage such situations efficiently. The main application domains, the recent research and the strengths and weaknesses of fuzzy clustering are presented in this article.

Keywords: machine learning, unsupervised learning, fuzzy clustering, customer relationship management, market segmentation

1. Introduction

Machine learning is the field of research devoted to the formal study of learning systems. This is a highly interdisciplinary field which borrows and builds upon ideas from statistics, computer science, engineering, cognitive science, optimization theory and many other disciplines of science and mathematics.

Cluster analysis belongs to the procedures of unsupervised machine learning, where no a-priori knowledge for the learning process is needed. The goal is to gain useful information about the structure of a given (complex) pattern set. Patterns (observations, data items, or feature vectors) are divided into *natural* groups (clusters). (*P. Perner, 2009*)

In this new era of uncertainty there is a clear need for data analysis and preprocessing. Advertising is growing more and more via social networks. As Brain Solis said: “Content is advertising”. For a 360^o view on the needs and desires of customers, marketers need a clear stratification of the information available on the internet, to know exactly what to include in their advertisements. For this reason clustering techniques are best for serving the marketers needs.

Fuzzy clustering is one method which can capture the uncertainty situation of real data and it is well known that fuzzy clustering can obtain a robust result as compared with conventional hard clustering.

This paper provides a clear presentation of the fuzzy clustering technique taking into consideration the unsupervised learning approach. The main applications and the recent research of the fuzzy clustering field are also being presented.

1.1 Motivation

A lot of study has been conducted for analyzing customer preferences in marketing. The main goal of customer profiling (or segmentation) is to build reliable customer models for targeted marketing campaigns; and consequently, a better profitability *(Romdhane L.B, 2010)*. Consequently, one can define data mining in customer profiling as being the technology that allows building customer models each describing the specific habits, needs, and behavior of a group of customers. Once discovered, these models can be used to classify new customers; and thereby, predict their special needs. The clustering techniques applicability in the service industries such as telecommunications, hotels, insurance, banking, retail, etc. represents a clear motivation, both for me and for the current managers, to use such techniques, in particular the fuzzy clustering technique. In real applications there is very often no sharp boundary between clusters so that fuzzy clustering is often better suited for the data. Membership degrees between zero and one are used in fuzzy clustering instead of crisp assignments of the data to clusters. The most prominent fuzzy clustering algorithm is the fuzzy c-means, a fuzzification of k-Means.

1.2 Theoretical Importance

Laviolette, Seaman, Barrett and Woodal have done us all a big favor by their well-written survey on fuzzy methods. Suppose we have n multivariate objects $x_i = (x_{i,1}, \ldots, x_{i,p})$ in p-dimensional space, and we would like to group these objects into k clusters. The aim is to obtain an assignment of objects to clusters, which can be described by a collection of nonnegative memberships:

$U_{it} \geq 0$ for all i= 1…n and t = 1…k, which have to satisfy the "partition constraint": $\sum u_{it} = 1$, t= 1..k, ¥ i = 1..n

This clearly implies that all $u_{it} \leq 1$. From here on, two different approaches are possible: If u_{it} is defined as a boolean (binary) variable, with the sole values 0 and 1, the clustering is called crisp (or hard). In this setting, object *i* belongs to exactly one cluster t (for which $u_{it} = 1$) and not to any other cluster t' (for which $u_{it'} = 0$). On the other hand, the framework of fuzzy clustering allows u_{it} to take on any value in the interval [0, 1]. This makes it possible to describe some ambiguities that often occur in real data, such as bridging objects and outliers.

An example with such ambiguities is shown in Figure 1. A crisp clustering method will assign object 6 to either the upper left cluster or to the upper right cluster, which is not very satisfactory because the object lies about equally far away from both. It is even harder to assign object 13 in a meaningful way because it lies between three main groups. A fuzzy clustering is better equipped to describe the situation. It can say that object 1 belongs mainly to the first cluster, whereas object 13 should be divided between all three clusters. (*Peter J. Rousseeuw, 1995*).

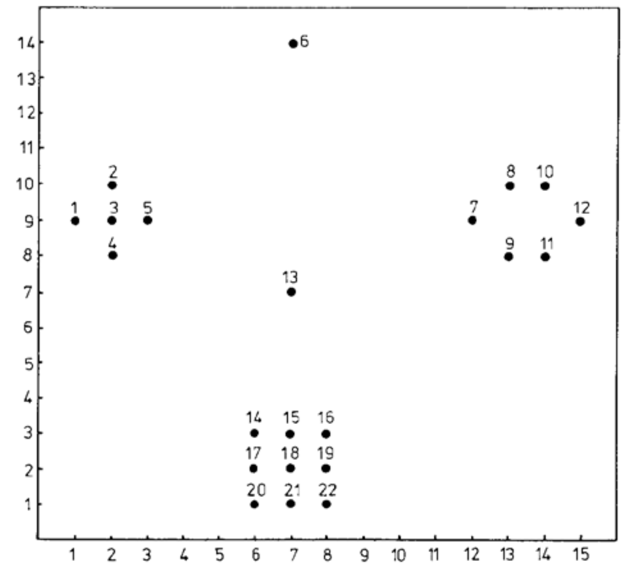


Figure 1. Data Set With Two Intermediate Objects.

1.3 Applications

Cluster analysis has been continuously developed and is used in many scientific disciplines such as biology, psychology, statistics, pattern recognition, economics and finance. (*P. Perner, 2009*)

Fuzzy clustering can be used to segment the market. The fuzzy clustering method can be used to modify a segmentation technique by generating a fuzzy score for each customer. This provides a more precise measure to the company in delivering value to the customer and profitability to the company. In the Wong's article, a case study was presented and two areas in a typical CRM model where the use of fuzzy theory can improve the decision making process were highlighted. The advantage of using fuzzy theory in CRM is that the business analyst can gain in-depth understanding into the data mining model. With the understanding of the model, the analyst can modify and add-on knowledge and experience into the model. Besides, fuzzy theory can handle uncertainties in the data more efficiently than traditional data mining techniques. (*Wong, 2001*)

In "*Robust Statistics and Fuzzy Industrial Clustering*" article a multidimensional fuzzy clustering analysis gives as a result a classification of sectors illustrating the different roles that each one plays in the economy. In the article is presented how fuzzy clustering and robust statistics should work together in this kind of studies, so the clustering can benefit from the use of robust statistics in data preparation, identification and computation of dissimilarities or deciding the best number of clusters and specially avoiding the dangerous effects coming from the presence of multivariate outliers. (*Diaz and Morillas, 2008*)

Other new applications can be found in geology (*Lucieer, 2009*), physics (*Wang and Pamela McCauley Bell, 1996*), biology, marketing research, educational research. (*Wikipedia*).

2. Unsupervised Cluster Analysis

2.1 What is unsupervised learning?

Consider a machine (or living organism) which receives some sequence of inputs $x_1,x_2,x_3,...$, where x_t is the sensory input at time t. This input, often called the data, could correspond to an image on the retina, the pixels in a camera, or a sound waveform. It could also correspond to less obviously sensory data, for example the words in a news story, or the list of items in a supermarket shopping basket.

In supervised learning the machine is also given a sequence of desired outputs $y_1,y_2,\ldots$, and the goal of the machine is to learn to produce the correct output given a new input. This output could be a class label (in classification) or a real number (in regression).

In unsupervised learning the machine simply receives inputs $x_1,x_2,...$, but obtains neither supervised target outputs, nor rewards from its environment. It may seem somewhat mysterious to imagine what the machine could possibly learn given that it doesn't get any feedback from its environment. However, it is possible to develop of formal framework for unsupervised learning based on the notion that the machine's goal is to build representations of the input that can be used for decision making, predicting future inputs, efficiently communicating the inputs to another machine, etc. In a sense, unsupervised learning can be thought of as finding patterns in the data above and beyond what would be considered pure unstructured noise. Two very simple classic examples of unsupervised learning are clustering and dimensionality reduction. (*Zoubin, 2004*)

2.2 Fuzzy Clustering

Conventional clustering means classifying the given observation as exclusive subsets (clusters). That is, it can be see clearly whether an object belongs to a cluster or not. However, such a partition is insufficient to represent many real situations. Therefore, a fuzzy clustering method is offered to construct clusters with uncertain boundaries, so this method allows that one object belongs to some overlapping clusters to some degree. In other words, the essence of fuzzy clustering is to consider not only the belonging status to the clusters, but also to consider to what degree do the objects belong to the clusters. (*Mika Sato-Ilic, Lakhmi C. Jain, 2006*)

Fuzzy clustering is a partition based clustering scheme and is particularly useful when there are no apparent clear groupings in the data set. Partitioning schemes provide automatic detection of cluster boundaries and in case of fuzzy clustering, these cluster boundaries overlap. Every individual data entity (a conformer, in this case) belongs to not one but all of the clusters with varying degrees of membership. However, there are very few instances where a partitioning scheme has been used to analyze families of molecular conformation.

It is also important to understand the difference between unsupervised classification and supervised classification. In supervised classification, a collection of *labeled* (preclassified) patterns is provided; the problem is to label a newly encountered, yet unlabeled, pattern. Typically, the given labeled (*training*) patterns are used to learn the descriptions of classes which in turn are used to label a new pattern. In the case of clustering, the problem is to group a given collection of unlabeled patterns into meaningful clusters. In a sense, labels are associated with clusters also, but these category labels are *data driven*; that is, they are obtained solely from the data. Clustering is useful in several exploratory pattern-analysis, grouping, decision- making, and machine-learning situations, including data mining, document retrieval, image segmentation, and pattern classification.

However, in many such problems, there is little prior information (e.g., statistical models) available about the data, and the decision-maker must make as few assumptions about the data as possible. It is under these restrictions that clustering methodology is particularly appropriate for the exploration of interrelationships among the data points to make an assessment (perhaps preliminary) of their structure. (*Zoubin, 2004*)

3. Recent Research

Fuzzy clustering algorithms were used for target selection in direct marketing problem. Target selection is the problem of finding groups of customers for a particular product in direct marketing. In the problem of direct marketing, manufacturing companies try to have a contact or maintain a direct relationship with customers in order to target them individually for particular product offers or maximizing the profit. The work of Ramathilagam and Kannan introduce fuzzy clustering methods and *k*-means for dividing the customer database into groups with similar properties called customer segments to maximize the profit or fund rising in direct marketing. (*Ramathilagam and Kannan, 2009*).

Applications of fuzzy clustering can also be found in medicine. In the paper of *Zdrenghea (2010)* is presented a way to use fuzzy clustering for generating fuzzy rule bases in the implementation of an intelligent agent that interacts with human for diagnosis establishment: The Medical Diagnostics System. The system is intended to be a software learning application mainly destined to orientate the resident doctors in the process of establishing a diagnostic for the patients they are examining.

In a recent paper of *Nanda (2010)*, a data mining approach for classification of stocks into clusters is presented. After classification, the stocks could be selected from these groups for building a portfolio. It meets the criterion of minimizing the risk by diversification of a portfolio. The clustering approach categorizes stocks on certain investment criteria.

Moreover, a new method for multi-variable fuzzy forecasting based on fuzzy clustering and fuzzy rule interpolation techniques, is presented in a paper by *Chen and Chang (2010).*

The research presented by *Trappey (2009)* et al in their article "*The analysis of customer service choices and promotion preferences using hierarchical clustering*" develops a computer-based customer preference segmentation method that automatically and distinctively matches the needs and wants of restaurant customers with promotional offers and discounts. The store manager uses hierarchical clustering to group the customers. For the first level of segmentation, customers are clustered based on frequency of visits and dining expenditures. Secondly, K-means clustering is used to analyze each sub-segment based on menu choice preferences. Given these results, the restaurant provides customized coupons and price discounts for each customer based on their previous preferences and behaviors. The study demonstrates an effective means to better manage and promote complex menu selections in a chain store or franchise restaurant environment. (*Trappey, 2009*)

Adem Karahoca and Dilek Karahoca (2011) proposed a method to prevent particular GSM operators' subscribers to pass other GSM operators. In the study, input features are clustered by x-means and fuzzy c-means clustering algorithms to put the subscribers into different discrete classes. Adaptive Neuro Fuzzy Inference System (ANFIS) is executed to develop a sensitive prediction model for churn management by using these classes. First prediction step starts with parallel Neuro fuzzy classifiers. After then, FIS takes Neuro fuzzy classifiers' outputs as input to make a decision about churners' activities.

Another recent research on Fuzzy Clustering Validity was done by *Weijin Chen, Huailin Dog, QingfenWu and Ling Lin (2010).*

A research concerning the common problems in fuzzy clustering algorithms was developed by *Min Min (2010).* In this study a combined fuzzy clustering based on F-Statistics was analyzed.

A "Novel Fuzzy Clustering-based Image Segmentation with Simultaneous Uneven Illumination Estimation" was presented by *Wei Wei-Yi, Li Zhan-Ming and Zhang Guo-Quan(2011)* in the article with the same name. A robust fuzzy clustering based segmentation method for noisy and intensity inhomogeneity caused by uneven illumination was discussed in the article.

In the article of *Hosein Hashemi, Abdolrahim Javaherian and Robert Babuska*: “A semi-supervised method to detect seismic random noise with fuzzy GK clustering”, a new method to detect random noise in seismic data using fuzzy Gustafson–Kessel (GK) clustering is presented. Using the knowledge of a human specialist together with the fuzzy unsupervised clustering, the method is a semi-supervised random noise detection. The efficiency of this method is investigated on synthetic and real seismic data for both pre- and post-stack data. The results show a significant improvement of the input noisy sections without harming the important amplitude and phase information of the original data. The procedure for finding the final weights of each clustered section should be carefully done in order to keep almost all the evident seismic amplitudes in the output section. The method interactively uses the knowledge of the seismic specialist in detecting the noise.

3.1 Fuzzy clustering validity indexes

Most validity indexes measure the degree of compactness and separation for the data structure in all of c clusters and then finds an optimal c that each one of these optimal c clusters is compact and separated from other clusters. If the data set contains some noisy points that may be far away from other clustered points, it can be visualized that validity indexes will take the noisy point into a compact and separated cluster. The validity function can help to validate whether it accurately presents the data structure or not, when the partition is obtained by a clustering method.(*Wu and Yang, 2005*)

Questions like “how many clusters are there in the data set?”, “Does the resulting clustering scheme fits our data set?”, “Is there a better partitioning for our data set?” call for clustering results validation(*Halkidi, Batistakis, Vazirgiannis, 2002*).

Many cluster validity indices offers conclusion that there is not generally the best validity index. Moreover, existing cluster validity indices are not very efficient in estimation of clusters of different sizes and densities and if are based only on fuzzy membership values are not efficient in validation of clusters of different sizes and densities.

Validity indices that are not based on average values are efficient in validation of clusters of different sizes and densities.

According to this I have found some new validity indexes that are promising a lot.

The Huang-index method

The results show that the proposed Huang-index method not only yields a superior clustering capability than the traditional clustering algorithm, but also yields a reliable classification and obtains a set of suitable decision rules extracted from the RS (rough set) theory. According to the article a method of cluster validity index that simultaneously provide the measurements of goodness of clustering on clustered data and of classification accuracy for complicated information systems based upon the PBMF-index method and rough set (RS) theory is developed. The maximum value of this index, called the Huang-index, not only provides the best partitioning, but also obtains the optimal accuracy of classification for the approximation sets. The traditional PBMF-index method is only used to ensure the formation of a small number of compact clusters with large separation between at least two clusters. In contrast to the traditional PBMF-index method, the Huang-index method extends the applications of unsupervised optimal cluster to the fields of classification. In the proposed algorithm, all the attributes of the data are first clustered into groups using the Fuzzy C-means (FCM) method. The clustered data are then used to identify approximate regions and classification accuracy and to calculate centroids of clusters for decision attribute based on the RS theory. Finally, all those calculated data are put into the proposed index method to find the cluster validity index. The validity of the proposed approach is demonstrated using the data derived from a hypothetical function of two independent variables and electronic stock data extracted from the financial database maintained by the Taiwan Economic Journal (TEJ).(*Kuang Yu Huang, 2010*)

Another important index was proposed by **Žalik** in a recent article. As he stated, cluster validity indices are used to validate results of clustering and to find a set of clusters that best fits natural partitions for given data set. Most of the previous validity indices have been considerably dependent on the number of data objects in clusters, on cluster centroids and on average values. They have a tendency to ignore small clusters and clusters with low density. Two cluster validity indices are proposed for efficient validation of partitions containing clusters that widely differ in sizes and densities. The first proposed index exploits a compactness measure and a separation measure, and the second index is based on an overlap measure and a separation measure. The compactness and the overlap measures are calculated from few data objects of a cluster while the separation measure uses all data objects. The compactness measure is calculated only from data objects of a cluster that are far enough away from the cluster centroids, while the overlap measure is calculated from data objects that are enough near to one or more other clusters. A good partition is expected to have low degree of overlap and a larger separation distance and compactness. The maximum value of the ratio of compactness to separation and the minimum value of the ratio of overlap to separation indicate the optimal partition. Testing of both proposed indices on some artificial and three well-known real data sets showed the effectiveness and reliability of the proposed indices. *(Žalik, 2010)*

Other important indexes from the literature are:

- **gath.geva**:. These indexes are only for the cmeans clustering algorithm
- **xie.beni**: This index is a function of the data set and the centroids of the clusters. Xie and Beni explained this index by writing it as a ratio of the total variation of the partition and the centroids and the separation of the centroids vectors. The minimum values of this index under comparison support the best partitions.
- **fukuyama.sugeno**: This index consists of the difference of two terms, the first combining the fuzziness in the membership matrix with the geometrical compactness of the representation of the data set via the prototypes, and the second the fuzziness in its row of the partition matrix with the distance from the i^{th} prototype to the grand mean of the data. The minimum values of this index also propose a good partition.
- **partition.coefficient**: An index which measures the fuzziness of the partition but without considering the data set itself. It is a heuristic measure since it has no connection to any property of the data. The maximum values of it imply a good partition in the meaning of a least fuzzy clustering.
- **partition.entropy**: It is a measure that provides information about the membership matrix without also considering the data itself. The minimum values imply a good partition in the meaning of a more crisp partition.
- **proportion.exponent**: It is a measure *P(U;k)* of fuzziness adept to detect structural variations in the partition matrix as it becomes more fuzzier. A crisp cluster in the partition matrix can drive it to infinity when the partition coefficient and the partition entropy are more sensitive to small changes when approaching a hard partition. Its evaluation does not also involve the data or the algorithm used to partition them and its maximum implies the optimal partition but without knowing what maximum is a statistically significant maximum.
- **separation.index (known as CS Index)**: This index identifies unique cluster structure with well-defined properties that depend on the data and a measure of distance. It answers the question if the clusters are compact and separated, but it rather seems computationally infeasible for big data sets since a distance matrix between all the data membership values has to be calculated. It also presupposes that a hard partition is derived from the fuzzy one.

4. Strong and Weak Points

One of the main advantages of fuzzy clustering is the ability to express ambiguity in the assignment of objects to clusters. But apart from this, experimental results prove that fuzzy clustering seems also to be more robust in terms of local minima of the objective function. As *F. Klawonn(2004)* stated in his article:

Fuzzy Clustering: Insights and a New Approach, the non-zero membership degree property of probabilistic clustering has a smoothing effect on undesired local minima of the objective function, and in the same time causes problems in fuzzy clustering. These bad effects can be avoided by the modified transformations replacing the fuzzifier (m). The disadvantage is that with a growing number of objects the amount of output of the results becomes immense, so that the information received often cannot be worked up.

Fuzzy clustering has a number of useful properties. For instance:

- It provides membership values which are useful for interpretation
- It is flexible with respect to distance used
- If some of the membership values are known this can be incorporated into the numerical optimisation.

(*Tormod Næs and Bjørn-Helge Mevik, 1999*)

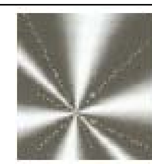
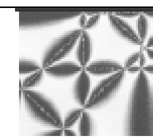
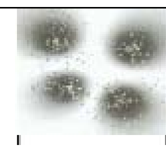
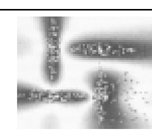
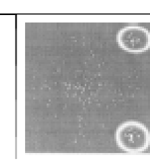
Another distinct advantage of fuzzy clustering over its crisp counterpart is that the continuous range of the u_{it} turns combinatorial functions into smooth functions. This makes it possible to design algorithms that are more likely to attain a global solution, whereas crisp techniques often wind up in the local solution. (*Peter J. Rousseeuw, 1995*).

The main difference between fuzzy clustering and other clustering techniques is that it generates fuzzy partitions of the data instead of hard partitions. Therefore, data patterns may belong to several clusters, having in each cluster different membership values.

Next, the advantages and disadvantages of different fuzzy clustering are described in Table 1. As shown in the table, fuzzy c means clustering method produces spherical clusters and it does not let a cluster change its shape dependent on the data. Whilst, Gustafson-Kessel proposed a method so a cluster could adapt to hyper-ellipsoidal shapes (*Martin, 2003*). Experiment by *Martin (2003)* shows that fuzzy c-means perform better when more number of clusters was used and Gustafson-Kessel algorithm performed best with relatively few clusters. The poorer results obtained by *Martin (2003)* by using Gustafson-Kessel was assumed to be caused by the clusters being limited to hyper-ellipsoidal shapes. Furthermore, this showed that clusters produced by Gustafson-Kessel did not supplement each other as very well as fuzzy cmeans clustering method. Hence, in among all fuzzy clustering method, fuzzy cmeans using Euclidean distance measures is better used in compound selection. (*Rozilawati Binti Dollah, 2006*)

With regards to performance, according to *Klawonn* the FCM needs to perform k (i.e. number of clusters) multiplications for each point, for each dimension (not counting also the exponentiation to take stiffness into account). This, plus the overhead needed for computing and managing the proximity vector, explains why FCM is quite slower than plain K-Means.(*Klawonn, 2004*)

Table 1. Different types of fuzzy clustering algorithm, Rozilawati Binti Dollah, 2006

	Advantages	**Disadvantages**	**Suitable for…**	**Shapes of clusters**
Fuzzy C Varieties (FCV)	Each cluster represents an r dimensional variety in the dimension of the data space	o The areas of high membership exceed beyond the line segments o A higher number of clusters increases the number of local minima	Lines, planes and hyper planes	
Adaptive Fuzzy Clustering(AFC)	Able to recognize elliptic or circular clusters	The eigenvalues have to be computed to update the prototypes, any changes are hardly visible	Line segments	
Fuzzy C Means (FCM) Algorithm	Few iterations steps already provide good approximation to the final solution	FCM tends to locate centroid in the neighborhood of the larger cluster and misses the small, well-separated cluster.	Spherical shape	
Gustafson Kessel (GK) Algorithm	Faster than AFC. In order to adapt to different structures in data, GK used the covariance matrix to capture ellipsoidal properties of clusters.	The clusters are narrower and the areas with higher membership are thinner	Line segments	
Gath-Geva (GG) Algorithm	Unlike FCM and GK algorithm, it is not based on objective function. It is a fuzzification of statistical estimators	Because the occurrence of the exponential function within the distance, the distance divided into two range, close and remote	Line segments	

4.1 Possible extensions

Uzay Kaymak and Magne Setnes (*2000*) in their article "Extended Fuzzy Clustering Algorithms" proposed two extensions in terms of the objective function. Objective function based fuzzy clustering algorithms such as the fuzzy c-means (FCM) algorithm have been used extensively for different tasks such as pattern recognition, data mining, image processing and fuzzy modeling. In order to obtain a good performance from a fuzzy clustering algorithm, a number of issues must be considered. These concern the shape and the volume of the clusters, the initialization of the clustering algorithm, the distribution of the data patterns and the number of clusters in the data. The proposed extensions are applied to Gustafson–Kessel and fuzzy c-means algorithms. The goal of this extension is to reduce the sensitivity of the clustering algorithms to bias in data distribution and to determine the number of clusters automatically.

Another extention of the fuzzy c-means algorithm was proposed by *Bernd Wiswedel and Michael R. Berthold (2007)*, which operates simultaneously on different feature spaces-so-called parallel universes- and also incorporates noise detection. The method assigns membership values of patterns to different universes, which are then adopted throughout the training. This leads to better clustering results since patterns not contributing to clustering in a universe are (completely or partially) ignored. The method also uses an auxiliary universe to capture patterns that do not contribute to any of the clusters in the real universes and therefore are likely to represent noise. The outcome of the algorithm is clusters distributed over different parallel universes, each modeling a particular, potentially overlapping subset of the data and a set of patterns detected as noise. One potential target application of the proposed method is biological data analysis where different descriptors for molecules are available but none of them by itself shows global satisfactory prediction results.

4.2 Hybrid Approaches

The fuzzy c-means algorithm is sensitive to initialization and is easily trapped in local optima. For this reason *Hesam Izakian and Ajuth Abraham (2010)* proposed a hybrid fuzzy clustering algorithm where the FCM algorithm is integrated with FPSO algorithm to form FCM-FPSO.

Another hybrid approach using fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS) was discussed by *Javier De Andreas (2010)* in his article "Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS)". In a first stage, clusters are created using fuzzy c-means. The clusters are classified into two groups: those that contain bankrupted companies and those that not. Then, a MARS model is created using such clusters as a part of the input information. The performance of the proposed model is better than those obtained with the following benchmark techniques: MARS, discriminant analysis and neural networks

5. Conclusions

In this paper it was presented the fuzzy clustering technique and the recent research in the field. Why do we want to choose a fuzzy clustering method and some advantages and disadvantages of fuzzy clustering were also presented. As a future work I will analyze the fuzzy c-means algorithm in comparison to other clustering algorithms to see how well will perform each in identifying customer preferences.

References

1. Adem Karahoca and Dilek Karahoca, GSM churn management by using fuzzy c-means clustering and adaptive neuro fuzzy inference system, Expert Systems with Applications Volume 38, Issue 3, Pages 1814-1822, March 2011

2. Bernd Wiswedel and Michael R. Berthold, Fuzzy Clustering in Parallel Universes, International Journal of Approximate Reasoning 45, 439-454, 2007

3. Chen Shyi-Ming and Yu-Chuan Chang, Multi-variable fuzzy forecasting based on fuzzy clustering and fuzzy rule interpolation techniques, Information Sciences, Volume 180, Issue 24, Pages 4772-4783, 15 December 2010

4. Diaz Barbara and Antonio Morillas, M. Nikravesh et al. (eds.), Robust Statistics and Fuzzy Industrial Clustering, Forging the New Frontiers: Fuzzy Pioneers II. 219 Springer-Verlag Berlin Heidelberg 2008

5. Halkidi Maria, Yannis Batistakis, Michalis Vazirgiannis, Clustering Validity Checking Methods: Part II, SIGMOD Record 31(3): 19-27, 2002

6. Hesam Izakian and Ajuth Abraham, Fuzzy Clustering Using Hybrid Fuzzy c-means and Fuzzy Particle Swarm Optimization, 2010

7. Javier De Andrés, Pedro Lorca, Francisco Javier de Cos Juez and Fernando Sánchez-Lasheras, Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS), Expert Systems with Applications, Volume 38, Issue 3, Pages 1866-1875, March 2011

8. Klawonn F., Fuzzy Clustering: Insights and a New Approach, Mathware & Soft Computing 11, p. 125-142, 2004

9. Kuang Yu Huang **,** Applications of an enhanced cluster validity index method based on the Fuzzy C-means and rough set theories to partition and classification**,** Expert Systems with Applications, Volume 37, Issue 12, Pages 8757-8769, December 2010

10. Lucieer V. and A. Lucieer, Fuzzy clustering for seafloor classification, Marine Geology, Volume 264, Issues 3-4, Pages 230-241, 15 August 2009

11. Martin, E. ,Pap-Smear Classification. Technical University of Denmark., 2003

12. Mika Sato-Ilic, Lakhmi C. Jain, Innovations in Fuzzy Clustering-Theory and Applications, Studies in Fuzziness and Soft Computing, Volume 205, Springer, 2006

13. Min Min, Study of Combined Fuzzy Clustering Algorithm Based on F-Statistics Hierarchy Clustering, Applied Mechanics and Materials, Vols 29-32, pp. 802-808, 2010

14. Nanda S.R., B. Mahanty And M.K. Tiwar**,** Clustering Indian Stock Market Data For Portfolio Management, Expert Systems With Applications, Volume 37, Issue 12, Pages 8793-8798, December 2010

15. Peter J. Rousseeuw, Technometrics, vol 37, no 3, august 1995

16. P. Perner, Machine Learning and Data Mining in Pattern Recognition, (Ed):MLDM 2009, LNAI 5632, pp. 337-346, 2009

17. Ramathilagam S. (Taiwan) And S. R. Kannan (India), Fuzzy Target Selection In Direct Marketing, Advances In Fuzzy Sets And Systems Volume 4, Issue 3, Pages 313 - 331 (October 2009)

18. Romdhane L.B., N. Fadhel, B. Ayeb , An efficient approach for building customer profiles from business data, Expert Systems with Applications 37 1573–1585, 2010

19. Rozilawati Binti Dollah, Aryati Binti Bakri, Mahadi Bin Bahari, Pm Dr. Naomie Binti Salim, Feasibility Study Of Fuzzy Clustering Techniques In Chemical Database For Compound Classification, 2006

20. Tormod Næs and Bjørn-Helge Mevik, The flexibility of fuzzy clustering illustrated by examples, Journal of Chemometrics ;13: 435–444, 1999

21. Trappey V. Charles, Trappey J.C. Amy, Ai-Che Chang, Ashley Y.L. Huang, The Analysis of

customer service choices and promotion preferences using hierarchical clustering, Journal of the Chinese Institute of Industrial Engineers, Vol. 26, No. 5, pp. 367-376, 2009

22. Žalik Krista Rizman and Borut Žalik**,** Validity index for clusters of different sizes and densities, Pattern Recognition Letters, Volume 43, Issue 10, Pages 3374-3390, October 2010

23. Zoubin Ghahramani, Unsupervised Learning, Advanced Lectures on Machine Learning LNAI 3174, 2004

24. Zdrenghea Vlad , Diana Ofelia Man, And Maria Tosa-Abrudan, Fuzzy Clustering In An Intelligent Agent For Diagnosis Establishment, Studia Univ. Babes - Bolyai, Informatica, Volume LV, Number 2, 2010

25. Uzay Kaymak, Magne Setnes, Extended Fuzzy Clustering Algorithms, ERIM Report Series Research in Management, November 2000

26. Weijin Chen, Huailin Dog, QingfenWu and Ling Lin, Research on Fuzzy Clustering Validity, Applied Mechanics and Materials, Vols. 40-41, pp. 174-182, 2011

27. Wei-Yi, W., L. Zhan-Ming and Z. Guo-Quan, Novel fuzzy clustering-based image segmentation with simultaneous uneven illumination estimation. Inform. Technol. J., 10: 607-610, 2011

28. Wu Kuo-Lung, Miin-Shen Yang, A cluster Validity index for fuzzy clustering, Pattern Recognition Letters 26,1275–1291, 2005