Semi-supervised Clustering in Fuzzy Rule Generation

Priscilla A. Lopes[1], **Heloisa A. Camargo**[1]

[1]Departamento de Computação – Universidade Federal de São Carlos (UFSCar)
Caixa Postal 676 – 13.565-905 – São Carlos – SP – Brazil

{priscilla_lopes,heloisa}@dc.ufscar.br

Abstract. *Inductive learning approaches traditionally categorized as supervised, which use labeled data sets, and unsupervised, which use unlabeled data sets in learning tasks. The great volume of available data and the cost involved in manual labeling has motivated the investigation of different solutions for machine learning tasks related to unlabeled data. The approach proposed here fits into this context: a semi-supervised clustering algorithm is applied to a partially labeled data set; the obtained results are used to automatically label the remaining data in the set; following, a supervised learning algorithm is used to generate fuzzy rules from the labeled data. The experiments show that this may be a promising solution for tasks that have encountered difficulties due to partially labeled data.*

1. Introduction

In recent years, the fields of machine learning and knowledge acquisition have been frequently focusing research related to Fuzzy Systems. These systems are based in the fuzzy sets theory and are mainly characterized by the use of linguistic variables in representing at least part of the knowledge. The utility of such systems is recognized in literature and still motivates investigations on methodologies to automatically build and optimize these systems aided by data sets.

Inductive learning distinguishes itself between supervised, making use of data with known classes (labeled data), and unsupervised learning, which utilizes data without classes (unlabeled data). Often enough, the scientific community addresses the need for research efforts to treat for the great volume of available data in the most diverse knowledge areas, due to the facility in obtaining this data accompanied by its high cost of manual interpretation and labeling. This situation refers to the study of learning mechanisms that consider both labeled and unlabeled data, which is known as semi-supervised learning [Klose and Kruse 2005, Zhu 2005].

Both techniques of supervised and unsupervised learning have been applied to fuzzy rule-based systems generation, the particular fuzzy systems of interest for this proposal. The current stage of research evidences that the combination of methodologies benefits the learning process as a whole, because it permits the balancing of advantages of complementary methods.

Among the supervised learning techniques, approaches that involve hybrid methods combining fuzzy set theory and other methodologies are highlighted, especially evolutionary algorithms [Cordón et al. 2004b] and neural networks [Jang et al. 1997].

As for unsupervised learning techniques, the analysis of clusters is a mechanism often explored for generation of fuzzy rules and sets from unlabeled data. The frequent approach is to apply a clustering algorithm and transform the result in a fuzzy rule base, in which each cluster defines one rule. The fuzzy sets in each dimension are obtained by projection of the clusters. Most times these projections are approximated to triangular and trapezoidal sets.

This work investigates the combination of semi-supervised clustering algorithms with the generation of fuzzy rules. The proposed approach applies a semi-supervised clustering algorithm to a partially labeled data set, using the results to automatically label the remaining data in the set. Later, a supervised learning algorithm is used to generate fuzzy rules from the labeled data.

2. Semi-supervised Data Clustering

Methods of supervised and unsupervised learning, although useful in determined applications, present disadvantages. In supervised learning there is the need of an already labeled data set for extraction of rules. Unsupervised learning depends on the selection of appropriate metric and distribution function, number of clusters to be created and appropriate representation of significant clusters in a data set.

With the increasing complexity of data sets in many domains, complete and manual data labeling becomes more difficult and onerous. In this scenario, the use both types of data, labeled and unlabeled, becomes more interesting. A growing number of publications and conferences about semi-supervised learning is observed, reporting successful applications, especially, in image processing [Grira et al. 2006, Pedrycz et al. 2008] and text classification [Liu and Huang 2003, Geng et al. 2009].

The publications suggest and analyze modifications to already known methods aiming to consider their application to a data set with little amount of labeled and mostly unlabeled data. The work of Zhu [Zhu 2005] presents some tendencies and characteristics of semi-supervised classification, such as self-training, co-training and generative models [Chapelle et al. 2006], and points towards other forms of semi-supervised learning, by means of clustering and regression tasks.

Typically, clustering methods do not use previous knowledge, such as class labels. Semi-supervised clustering algorithms on other hand include mechanisms to consider previous information in the process of generating clusters. These mechanisms include: modification of the objective function, as to include restrictions satisfaction [Pedrycz and Waletzky 1997]; reinforcement of restrictions during the clustering process [Wagstaff et al. 2001, Grira et al. 2005]; initialization and restriction of clustering based on labeled examples [Bensaid et al. 1996, Basu et al. 2002].

The clustering methods that incorporate semi-supervision may be divided in two approaches, depending of available knowledge: seed-based approach [Bensaid et al. 1996, Pedrycz and Waletzky 1997, Basu et al. 2002] and pair-restriction-based approach [Wagstaff et al. 2001, Basu et al. 2004, Grira et al. 2008]. Seeds are labeled examples within a data set that may be used to establish restrictions to the algorithm, restrictions between pairs of examples and to define cluster labels. The pair-wise restrictions may be of type *must-link*, indicating that a pair of examples must belong to the same cluster, or *cannot-link*, indicating that a pair of examples must belong to distinct clusters.

An example of semi-supervised clustering is the semi-supervised Fuzzy C-Means algorithm (ssFCM) [Bensaid et al. 1996]. This method is an adaptation of the Fuzzy C-Means (FCM) algorithm, that considers information in the form of seeds to improve clustering performance.

The suggested modifications were made aiming to solve three problems from the application of original FCM. The first would be the difficulty in selecting the c number of clusters, due to lack of knowledge about the data set. The second is the problem of defining appropriate labels to each cluster after the clustering process. The last problem relates to objective functions that tend to equal the number of examples in each cluster.

The ssFCM algorithm tries to minimize the described problems. Firstly, the initial FCM data set would be substituted by a union of the labeled data, X^l, and unlabeled data, X^u. Let n be the number of labels represented in X^l, $c = n$. The definition of fuzzy pseudo-partitions is changed so that the membership of labeled examples is defined as 1 and is not altered during the updates of the pseudo-partition. Lastly, a weight is added to each labeled example (w_k) to calculate cluster prototypes, which is defined according to the degree of influence of each example.

3. Semi-supervised Clustering in Fuzzy Rule Generation

The goal of this work is to explore the use of clustering algorithms in the generation of fuzzy rules. This section initiates with a brief description of the most used mechanisms in fuzzy rule generation. Following, there is a discussion on the role of clustering methods in this process, mainly, by the unsupervised version.

3.1. Fuzzy Rule-based Systems Generation

Fuzzy Rule-based Systems (FRBS) are composed by two main components: the Knowledge Base (KB) and the Inference Mechanism (IM). The automatic generation of these systems is a frequent topic in recent researches and characterizes the approximation between the areas of fuzzy modeling and machine learning [Hullermeier 2005]. As the KB of a FRBS includes a Data Base (DB), which contains the fuzzy sets, and a Rule Base (RB), which contains the rules, the methods for generation of these systems vary from adjustment of fuzzy sets, going through generation of one of the components of the KB, to the simultaneous generation of all the components, including parameters of the IM.

Many approaches have been used to automatically generate the KB from numeric data that represent samples or examples of the problem. Among the most successful techniques are the clustering algorithms [Liao et al. 2003], the gradient-based methods [Nomura et al. 1992], the fuzzy decision trees [Janikow 1998], the neural networks [Jang et al. 1997] and the genetic algorithms [Cordón et al. 2004a].

The use of clustering aiming to build a FRBS consists in identifying regions in an input space that may form the antecedent of a rule. The results of the fuzzy clustering may be transformed in a fuzzy RB such that each cluster represents a fuzzy rule. This approach is discussed in the next section.

The hybrid approaches, which combine different methodologies, specially neural networks and evolutionary computing, with fuzzy sets theory, are particularly important for fuzzy rule learning. Evolutionary algorithms in general and genetic algorithms in particular have been extensively used for FRBS learning. In this approach there is a great variety of mechanisms to optimize or build a FRBS or its parts, sequentially, simultaneously or even as a combination of other methods.

Fuzzy Systems combined with a learning process based on genetic algorithms are named Genetic Fuzzy Systems (GFS). The research on genetic rule-base fuzzy systems has already originated an expressive number of work and continues in activity nowadays. Important contributions that characterize the state of the art and the main tendencies may be found in the many special editions on the subject, such as [Cordón et al. 2004b, Casillas et al. 2007, Herrera 2008, Casillas and Carse 2009].

The GFS are particularly relevant to this work, as long as some of the supervised methods for rule generation utilized in this work were chosen among the ones that fit this GFS proposal.

3.2. Fuzzy Clustering in Generating Fuzzy Rule-based Systems

The most common approach to generate rules by fuzzy clustering is to consider each cluster as a possible rule. Thus, after finding clusters through an algorithm, the antecedent of the rule is determined by projecting the cluster on each of the input space dimensions to obtain propositions in the form "V is A". The conjunction of these propositions forms the antecedent of the rule and each cluster is associated to a class. The disadvantage of this approach is that each rule makes use of its own fuzzy sets, generated by the projection, which may cause linguistic interpretability problems [Klose and Kruse 2005]. Examples of pioneer work following this approach are [Sugeno and Yasukawa 1993, Yager 1993, Babuska et al. 1994].

In recent work it is frequently found approaches that combine clustering methods with other learning mechanisms for the generation of fuzzy rules. In [Juang 2005], for instance, the proposed method combines on-line clustering and genetic algorithm in generating FRBS, with reinforcement. Other approaches that combine genetic algorithms with clustering may be found in [Saez et al. 2008]. In [Lee et al. 2008] a new iterative fuzzy clustering algorithm is presented. The fuzzy rules are obtained by an iterative process of selecting clusters with supervision based on notions of purity and separability of clusters.

Finally, it must be mentioned that clustering methods are very utilized in the generation of fuzzy sets, combined with the fuzzy rule generation process made by another method [Liao et al. 2003].

The proposal developed here concerns the combination of clustering methods with fuzzy systems generation in a way that is different from the ones usually encountered in the literature. The proposal is presented in the following section.

4. Proposed Approach

The goal of the proposed approach is to explore and evaluate the use of semi-supervised clustering methods in generating fuzzy rule bases. The main idea is to automatically label partially labeled data sets, based on the partition obtained by the application of semi-supervised clustering methods and, then, generate a fuzzy rule-based system with the aid of a supervised classification method.

The main motivation for this work was to investigate a hybrid mechanism that use partially labeled data, by the combination of two complementary methodologies. While the generation of fuzzy rules is usually done by mechanisms that require labeled data, the increasing availability of unlabeled data evidences the need to research methods that are adequate for such situations. On the other hand, clustering methods, traditionally applied to unlabeled data, present difficulties relative to validation and interpretation of groups resulting from the partitioning, which might not favor its utilization to generate rules. The proposed method aims to study alternatives to surpass those difficulties.

The method developed in this work is divided in two steps: (1) clustering and labeling; (2) rule generation . These steps are described in the following sections.

4.1. Clustering and Labeling Step

This step aims to automatically label the training data, according to the resulting partition of the semi-supervised algorithm. The labeled data can then be used in supervised learning, specifically, in the generation of fuzzy rules.

The clustering was based on the seeded-based ssFCM algorithm. The FCM algorithm was used for comparison and clustering validation. The comparisons between the clustering algorithms were based on the heterogeneity index R [Carvalho et al. 2006].

4.2. Rule Generation Step

The rule generation step consists of the application of a supervised fuzzy rule generation method on the labeled data set resulting from the previous step. Aiming at providing enough results to support the comparative analysis, four different rule generation algorithms have been used in this work.

One of the algorithms used is a non-evolutionary fuzzy rule base generation method, Fuzzy Rule Learning Model [Chi et al. 1996, Cordón et al. 1999]. The Fuzzy Rule Learning Model (CRW) builds a fuzzy rule base by means of a technique adapted from Wang & Mendel method, also considering weights that improve the classification. The other algorithms selected for the rule generation task are evolutionary methods. The hybrid fuzzy genetics-based machine learning algorithm (IH) [Ishibuchi et al. 2005] is a method that implements a hybridism between two famous approaches: Michigan and Pittsburgh. The steady-state algorithm for extracting fuzzy classification rules from data (SG)[Mansoori et al. 2008] is a genetic algorithm with a finite number of generations, which is related to the dimension of the evaluated problem. The goal of the SG is to extract a more compact and legible rule base. The structural learning algorithm on vague environment (SL) [González and Perez 2001] method is a genetic algorithm based on the iterative approach.

The selected methods were applied using the KEEL (Knowledge Extraction based on Evolutionary Learning) tool [Alcala-Fdez et al. 2009]. The evaluation of obtained results were based on the error rate, i.e., the percentage of incorrectly classified test instances.

The main goal of the experiments is to compare the results of the rule bases generated from automatically labeled data set and originally labeled data set.

Table 1. Data sets used in experiments

Data Set	Instances	Attributes	Classes
Apendicitis	109	9	2
Balance	625	4	3
Bupa	345	6	2
Ecoli	336	7	8
Glass	214	9	7
Haberman	306	3	2
Ionosphere	351	33	2
Iris	150	4	3
Monk 2	432	6	2
New Thyroid	215	5	3
Pima	768	8	2
Sonar	208	60	2
Spambase	4597	57	2
SPECTF Heart	267	44	2
Texture	5500	40	11
Vehicle	846	18	4
WDBC	569	30	2
Wine	178	13	3
Yeast	1484	8	10

5. Experiments and Results

Thirteen popular data sets, shown in Table 1were used to validate the proposal and evaluate the results obtained by the clustering and classification methods. The numerical data sets used are available at the UCI repository for machine learning [Frank and Asuncion 2010]. Cross-validation for 5 folds was adopted to compare the performance of clustering and classification algorithms.

The goals of the proposal allow the validation to be divided in two parts: the validation of results obtained by clustering methods and the validation of the fuzzy rule-base for classification generated by the rule base generation methods. Specifications of each part are described is sections 5.1 and 5.2.

5.1. Clustering Validation

The validation of clustering methods is given by an index obtained from the results of experiments considering two different fuzzy clustering algorithms:

FCM Unsupervised method. The original labels were removed from the data set. The metric for instance dissimilarity metric was the Euclidean distance.

ssFCM Semi-supervised method, based on seeds. The original labels of the training sets were partially removed and the remaining labels were used as seeds. Experiments were made with sets of seeds sized at approximately 10%, 20% and 30% of instances. The selection of seeds happened in two different ways: random selection of n instances per label and random selection of instances from the complete training set.

The implementation of the clustering algorithms uses a few structures of the WEKA [Hall et al. 2009] data mining software package.

This work utilizes an index based on sum of squares (SSQ) to evaluate the fuzzy clustering results applied to quantitative data sets. The overall SSQ and between-cluster SSQ metrics [Carvalho et al. 2006] are used in the definition of the interpretation index.

The overall SSQ metric (T) evaluates the general heterogeneity for all n examples in the data set, according to the distance function used in the clustering algorithm. This metric relates each example of the data set with the general data set prototype.

Between-cluster SSQ metric (B) evaluates the dispersion of cluster prototypes and, thus, the difference among all the clusters obtained with the application of the clustering algorithm. It relates the cluster prototypes with the general data set prototype and uses cluster membership degrees obtained by the clustering.

The general heterogeneity index (R) is obtained by $R = \frac{B}{T}$ were high values of R indicate clusters that are more homogeneous and better represented by the cluster prototype.

5.2. Classification Validation

The validation of the fuzzy rule-base generation is given by comparing and evaluating results obtained by the application of classification algorithms to the data sets with the original labels and on the data sets that have gone through the semi-supervised clustering and labeling.

Comparison of the results is given by the Friedman test [Demšar 2006]. A post-hoc test is applied when significant difference between results is verified.

Generally, the Friedman test is applied to evaluate algorithms. In this work, the goal is to verify whether the application of classification algorithms to data sets that were automatically labeled after being clustered by a semi-supervised algorithm influences the performance of the classification algorithm.

The results for the Friedman test and the post-hoc were given by the data analyses software GraphPad (http://www.graphpad.com/).

5.3. Results

The goal of a clustering algorithm is finding the optimal partition of a determined data set. This implies in partitioning the data as to obtain groups that are most homogeneous and compact as possible, given a proximity measure.

One of the difficulties in clustering tasks is that sets with high dispersion of data may influence the results of the clustering in a negative way. Table 2 indicate the R index, as presented in section 5.1, for results on the clustering of the used data sets. The applied clustering methods were FCM and ssFCM. In Table 2, <data-set-name>$_{(a,b,c)}$, a, b and c are the number of groups in which the data was partitioned and ssFCM$_i$, i is the percentage of seeds relative to the number of total instances in the training data set

Low values for R indicate a considerably disperse set. It is noted, in Table 2, that, as the number of groups in FCM clustering increases, the value R also increases. Data partitioning using higher numbers of groups may generate better values for the R index, although that does not always indicate good partitioning. The value of R, for some data sets, is inversely proportioned to the number of initial seeds. The disperse characteristic

Table 2. General heterogeneity index (R) for FCM and ssFCM

Data Set	FCM$_a$	FCM$_b$	FCM$_c$	ssFCM$_{10\%}$	ssFCM$_{20\%}$	ssFCM$_{30\%}$
Balance$_{(2,3,4)}$	0.0042	0.0366	0.0818	0.0192	0.0168	0.0121
Bupa$_{(2,3,4)}$	0.0001	0.0187	0.0717	0.0000	0.0000	0.0000
Haberman$_{(2,3,4)}$	0.8935	0.9432	0.9607	0.9049	0.8778	0.8628
Ionosphere$_{(2,3,4)}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Iris$_{(2,3,4)}$	0.5224	0.6476	0.7188	0.6800	0.6815	0.6825
Monk2$_{(2,3,4)}$	0.0174	0.0436	0.0697	0.0221	0.0114	0.0070
New Thyroid$_{(2,3,4)}$	0.6579	0.7499	0.8184	0.6985	0.6842	0.6931
Pima$_{(2,3,4)}$	0.0000	0.0006	0.0072	0.0000	0.0000	0.0000
Sonar$_{(2,3,4)}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
SPECTF Heart$_{(2,3,4)}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Vehicle$_{(3,4,5)}$	0.0026	0.0041	0.0058	0.0019	0.0012	0.0005
WDBC$_{(2,3,4)}$	0.0025	0.0056	0.0161	0.0020	0.0016	0.0012
Wine$_{(2,3,4)}$	0.0038	0.0116	0.1196	0.0003	0.0003	0.0003

Table 3. Error percentage for rule-base (O)

Data Set	CRW	IH	SG	SL
Balance	9.30	15.80	25.00	19.00
Bupa	40.87	29.57	39.89	36.20
Haberman	26.23	24.80	25.86	26.06
Ionosphere	18.61	34.61	23.29	5.41
Iris	7.17	4.83	6.08	3.17
Monk 2	30.57	4.09	19.45	2.78
New Thyroid	14.48	5.47	10.23	8.20
Pima	26.01	23.5	28.24	24.78
Sonar	20.41	51.05	27.14	21.27
SPECTF Heart	18.53	26.97	21.11	20.13
Vehicle	36.61	47.50	48.11	28.53
WDBC	5.59	8.22	8.19	7.95
Wine	4.30	8.87	9.19	4.38

of data sets may influence the partitioning using ssFCM, "confusing" the algorithm as the number of seeds increases.

Tables 3 and 4 present the error rate for experiments made with algorithms CRW, IH, SG and SL, with the different types of data. Table 3 contains the results for the rule-base generated from the 100% original labeled data set (O). Table 4 presents the results obtained from data set with planned seeds (P) and random seeds (R), which went through a labeling process after being clustered by the ssFCM algorithm.

An initial evaluation of Tables 3 and 4 suggest that the difference in results for the 3 types of data set (original and automatically labeled) are not significant.

In the first evaluation strategy, the Friedman test was among the rule-base generation method for each data set. The goal was to verify if an algorithm a_1 that performs significantly better than another algorithm a_2, when applied to data set O, will still perform better when applied to data sets P and R. The results obtained confirm the initial evaluation of the error rates table: the results are not significantly different.

In the second evaluation strategy used to analyze the results of the experiments, the comparison was made among the different data types for each rule generation method. Considering that the labeling based on the semi-supervised clustering is successful when

Table 4. Error percentage for rule-base (P and R)

Data Set	$\mathbf{CRW}_P$	$\mathbf{IH}_P$	$\mathbf{SG}_P$	$\mathbf{SL}_P$	$\mathbf{CRW}_R$	$\mathbf{IH}_R$	$\mathbf{SG}_R$	$\mathbf{SL}_R$
Balance	14.44	15.48	15.00	15.18	20.68	19.68	22.14	17.04
Bupa	19.13	11.01	14.89	9.06	19.82	13.74	17.43	10.66
Haberman	32.21	9.88	35.95	16.9	29.74	11.15	70.32	12.25
Ionosphere	25.71	38.50	28.34	16.24	27.60	46.09	28.34	15.33
Iris	8.50	6.17	7.42	6.50	8.67	6.75	8.25	7.09
Monk 2	46.93	11.48	16.42	8.94	45.86	8.51	10.20	7.58
New Thyroid	41.16	41.16	41.16	41.16	22.85	11.63	21.57	18.67
Pima	22.62	14.06	21.26	17.97	21.78	12.50	20.93	15.02
Sonar	25.72	52.76	28.59	22.71	25.73	52.76	28.60	22.71
SPECTF Heart	18.53	26.97	21.11	20.13	16.20	31.07	28.33	17.75
Vehicle	18.91	22.31	15.82	15.38	20.97	23.52	20.91	16.11
WDBC	7.58	8.76	4.55	3.62	7.58	6.69	4.29	3.15
Wine	4.30	8.87	9.19	4.38	23.36	19.22	16.66	14.18

the labels are as close as possible to the original ones, that goal is reached if the performance differences between the algorithms is not significant.

As this second application of the Friedman test showed no significant differences, it is inferred, according to earlier considerations, that the resulting labels obtained after the semi-supervised clustering may be as good as the ones in the original data sets.

6. Conclusion

This work presents a semi-supervised learning proposal that utilizes a semi-supervised clustering approach to label data for posterior application of a method to generate a fuzzy rule-base for classification.

This proposal suggests a solution for machine learning problems in domains that are characterized by great volume of data and varied data types, but have a relatively small set of labeled instances.

The results of this work suggest that considering semi-supervised clustering in labeling data may be a good approach for learning tasks in those contexts. The discussion presented here confirm the relevance and importance of the proposed work.

Relevant questions and considerations were raised due to the research involved in this proposal, which may be explored in future work: adaptation and proposal of semi-supervised clustering methods, aiming partially classified data labeling; comparison of results between distinct semi-supervised learning approaches, evaluation of results obtained by application of the proposed method as a solution for problems in specific domains.

Acknowledgment

The authors would like to thank E-Biz solution and CAPES for their support.

References

Alcala-Fdez, J., Sánchez, L., García, S., del Jesus, M., Ventura, S., Garrell, J., Otero, J., Romero, C., Bacardit, J., Rivas, V., Fernández, J., and Herrera, F. (2009). Keel: A software tool to assess evolutionary algorithms for data mining problems. *Soft Computing*, 13(3):307–318.

Babuska, R., Jager, R., and Verbruggen, H. (1994). Interpolation issues in sugeno-takagi reasoning. In *Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the Third IEEE Conference on*, pages 859–863.

Basu, S., Banerjee, A., and Mooney, R. J. (2002). Semi-supervised clustering by seeding. In *ICML '02: Proceedings of the Nineteenth International Conference on Machine Learning*, pages 27–34.

Basu, S., Banerjee, A., and Mooney, R. J. (2004). Active semi-supervision for pairwise constrained clustering. In *In Proceedings of the 2004 SIAM International Conference on Data Mining*, pages 333–344.

Bensaid, A. M., Hall, L. O., Bezdek, J. C., and Clarke, L. P. (1996). Partially supervised clustering for image segmentation. *Pattern Recognition*, 29(5):859–871.

Carvalho, F. A. T., Tenorio, C. P., and Junior, N. L. C. (2006). Partitional fuzzy clustering methods based on adaptive quadratic distances. *Fuzzy Sets and Systems*, 157(21):2833–2857.

Casillas, J. and Carse, B. (2009). Special issue on "genetic fuzzy systems: Recent developments and future directions". *Soft Comput.*, 13(5):417–418.

Casillas, J., Herrera, F., Pérez, R., del Jesús, M. J., and Villar, P. (2007). Special issue on genetic fuzzy systems and the interpretability-accuracy trade-off. *Int. J. Approx. Reasoning*, 44(1):1–3.

Chapelle, O., Schölkopf, B., and Zien, A. (2006). *Semi-Supervised Learning*. MIT Press, Cambridge, MA.

Chi, Z., Yan, H., and Pham, T. (1996). *Fuzzy Algorithms: With Applications To Image Processing and Pattern Recognition*. World Scientific.

Cordón, O., del Jesus, M., and Herrera, F. (1999). A proposal on reasoning methods in fuzzy rule-based classification systems. *International Journal of Approximate*, 20(1):21–45.

Cordón, O., Gomide, F., Herrera, F., Hoffmann, F., and Magdalena, L. (2004a). Special issue on genetic fuzzy systems. *Fuzzy Sets and Systems*, 141.

Cordón, O., Gomide, F., Herrera, F., Hoffmann, F., and Magdalena, L. (2004b). Ten years of genetic fuzzy systems: current framework and new trends. *Fuzzy Sets and Systems*, 141(1):5–31.

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.*, 7:1–30.

Frank, A. and Asuncion, A. (2010). UCI machine learning repository. http://archive.ics.uci.edu/ml.

Geng, C., Yuquan, Z., Jianing, T., and Tianhan, H. (2009). An algorithm of semi-supervised web-page classification based on fuzzy clustering. In *Information Technology and Applications, 2009. IFITA '09. International Forum on*, volume 1, pages 3–7.

González, A. and Perez, R. (2001). Selection of relevant features in a fuzzy genetic learning algorithm. *IEEE Transactions on Systems and Man and and Cybernetics and Part B: Cybernetics*, 31(3):417–425.

Grira, N., Crucianu, M., and Boujemaa, N. (2005). Active semi-supervised fuzzy clustering for image database categorization. In *MIR '05: Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval*, pages 9–16. ACM.

Grira, N., Crucianu, M., and Boujemaa, N. (2006). Fuzzy clustering with pairwise constraints for knowledge-driven image categorisation. *Vision, Image and Signal Processing, IEEE Proceedings -*, 153(3):299–304.

Grira, N., Crucianu, M., and Boujemaa, N. (2008). Active semi-supervised fuzzy clustering. *Pattern Recognition*, 41(5):1851–1861.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The weka data mining software: an update. *SIGKDD Explor. Newsl.*, 11(1):10–18.

Herrera, F. (2008). Genetic fuzzy systems: Taxonomy and current research trends and prospects. *Evolutionary Intelligence*, 1:27–46.

Hullermeier, E. (2005). Experience-based decision making: A satisficing decision tree approach. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 35(5):641–653.

Ishibuchi, H., Yamamoto, T., and Nakashima, T. (2005). Hybridization of fuzzy gbml approaches for pattern classification problems. *IEEE Transactions on Systems and Man and Cybernetics - Part B: Cybernetics*, 35(2):359–365.

Jang, J. S. R., Sun, C. T., and Mizutani, E. (1997). *Neuro-Fuzzy and Soft Computing*. Matlab Curriculum Series. Prentice Hall.

Janikow, C. Z. (1998). Fuzzy decision trees: issues and methods. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 28(1):1–14.

Juang, C.-F. (2005). Combination of online clustering and q-value based ga for reinforcement fuzzy system design. *Fuzzy Systems, IEEE Transactions on*, 13(3):289–302.

Klose, A. and Kruse, R. (2005). Semi-supervised learning in knowledge discovery. *Fuzzy Sets and Systems*, 149:209–233.

Lee, H. E., Park, K. H., and Bien, Z. Z. (2008). Iterative fuzzy clustering algorithm with supervision to construct probabilistic fuzzy rule base from numerical data. *IEEE T. Fuzzy Systems*, 16(1):263–277.

Liao, T. W., Celmins, A. K., and II, R. J. H. (2003). A fuzzy c-means variant for the generation of fuzzy term sets. *Fuzzy Sets and Systems*, 135(2):241–257.

Liu, H. and Huang, S. T. (2003). A genetic semi-supervised fuzzy clustering approach to text classification. In *In Proceedings of the 4th International Conference on Advances in Web-Age Information Management (WAIMâ03), Lecture Notes in Computer Science*, pages 173–180.

Mansoori, E., Zolghadri, M., and Katebi, S. (2008). Sgerd: A steady-state genetic algorithm for extracting fuzzy classification rules from data. *IEEE Transactions on Fuzzy Systems*, 16(4):1061–1071.

Nomura, H., Hayashi, I., and Wakami, N. (1992). A learning method of fuzzy inference rules by descent method. In *Proc. IEEE Int Fuzzy Systems Conf*, pages 203–210.

Pedrycz, W., Amato, A., Di Lecce, V., and Piuri, V. (2008). Fuzzy clustering with partial supervision in organization and classification of digital images. 16(4):1008–1026.

Pedrycz, W. and Waletzky, J. (1997). Fuzzy clustering with partial supervision. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 27(5):787–795.

Saez, Y., Zazo, F., and Isasi, P. (2008). A study of the effects of clustering and local search on radio network design: Evolutionary computation approaches. In *Hybrid Intelligent Systems, 2008. HIS '08. Eighth International Conference on*, pages 951–954.

Sugeno, M. and Yasukawa, T. (1993). A fuzzy-logic-based approach to qualitative modeling. *Fuzzy Systems, IEEE Transactions on*, 1(1):7.

Wagstaff, K., Cardie, C., Rogers, S., and Schrödl, S. (2001). Constrained k-means clustering with background knowledge. In *ICML '01: Proceedings of the Eighteenth International Conference on Machine Learning*, pages 577–584. Morgan Kaufmann Publishers Inc.

Yager, R. (1993). On a hierarchical structure for fuzzy modeling and control. *Systems, Man and Cybernetics, IEEE Transactions on*, 23:1189–1197.

Zhu, X. (2005). Semi-supervised learning literature survey. Technical Report 1530, Computer Sciences, University of Wisconsin-Madison.