Evolutionary Induction of Descriptive Fuzzy Rules in a Market Problem

María José del Jesus, Pedro González
Department of Computer Science
University of Jaén
Jaén E-23071, Spain
E-mail: mjjesus,pglez@ujaen.es

Francisco Herrera
Department of Computer Science
and Artificial Intelligence
University of Granada
Granada E-18071, Spain
E-mail: casillas@decsai.ugr.es

Mikel Mesonero
Department of Organisation
and Marketing
University of Mondragón
Oñati (Gipuzkoa) E-20560, Spain
E.mail: mmesoner@eteo.muni.es

Abstract—**Nowadays, face to face contact with the client is fundamental to the development of marketing acts. In this sense trade fairs are a basic instrument in company marketing policies, especially in Industrial Marketing. In this paper, we study the use of Soft Computing methodologies, specifically Fuzzy Logic and Genetic Algorithms, in the design of the Data Mining algorithms most proper to this problem. We present an evolutionary model for the descriptive induction of rules which describe subgroups, including a genetic algorithm in an iterative model to extract a variable number of fuzzy or crisp rules. The knowledge discovered with our proposal for each value of the target variable is interesting, understandable and have a high confidence and an adecuate support.**

I. Introduction

In the field of business an interesting problem is the study of the influence that the planning variables of a trade fair have over the successful achievement of its objectives. Faced with a real problem of this type a data mining algorithm should extract relevant interesting information concerning each of the effectiveness groups by which the stands have been arranged. The information obtained must be open to interpretation so as to be useful for the policies of trade fair planning. This problem is approached in this paper by a Genetic Fuzzy System included in the area of Knowledge Discovery in Databases (KDD).

KDD is wide ranging process defined as the non trivial process of identifying valid, original, potentially useful patterns from data [8] and covers distinct stages: the comprehension of the problem, the comprehension of the data, pre-processing of the data, data mining and post-processing (assessment and interpretation of the models). The data mining stage is responsible for automatic knowledge discovery of a high level and from information obtained from real data.

A data mining algorithm can discover knowledge using different representation models and techniques from two different perspectives:

- Predictive induction, whose objective is the discovery of knowledge for classification or prediction.
- Descriptive induction, whose fundamental objective is the discovery of interesting knowledge from the data.

Considering the characteristics of the problem to be solved, the obtention of simple rules which provide conclusive information about the efficiency of the stands in trade fairs, the most suitable approach is descriptive induction.

A subdivision of descriptive induction algorithms which has recently received a lot of attention from researchers is subgroup discovery which, given a set of data and having a property of interest to the user, attempts to locate subgroups which are statistically “most interesting”. The concept was initially formulated by Klosgen in his rule learning algorithm EXPLORA [18], and by Wrobel in the algorithm MIDOS [28]. The SD algorithm [10] or the CN2-SD algorithm [21] have been proposed later. This proposals are adaptations of classification rule extraction models for the subgroup discovery task, but currently interest is starting to be shown in the development of subgroup discovery approaches by modifying association rule extraction algorithms [20].

In this paper the subgroup discovery problem is approached with a Genetic Fuzzy System which hybridise the approximate reasoning method of fuzzy systems with the learning capabilities of Genetic Algoritms [5]. The proposal is an evolutionary model for the induction of descriptive fuzzy or crisp rules which describe subgroups. It includes a genetic algorithm in an iterative model which extracts rules when some examples are left uncovered, and the rules obtained surpass a given confidence level which is specified by the user.

To do so, the paper is arranged in the following way: In Section 2, the market problem and the kind of knowledge the user is interested in are dealt with. In section 3 the genetic approaches proposed in the bibliography to induce descriptive rules are outlined. The use of Fuzzy Logic in this kind of algorithm is described in Section 4. The evolutionary approach to obtain descriptive fuzzy rules is explained in Section 5, and in Section 6 the experimentation carried out and the analysis of results are explained. Finally, the conclusions and further research are outlined.

II. The Extraction of Useful Information on Trade Fairs

This study deals with a market problem analysed in the Department of Organisation and Marketing of the University of Mondragón, Spain: the extraction of useful information on trade fairs [22].

Businesses consider trade fairs to be an instrument which facilitates the attainment of commercial objectives such as contact with current clients, the securing of new clients, the taking of orders, and the improvement of the company image amongst others [13]. One of the main inconveniences in this type of trade fair is the elevated investment which they imply in terms of both time and money. This investment sometimes coincides with a lack of planning which emphasises the impression that trade fairs are no more than an "expense" which a business must accept for various reasons such as tradition, client demands, and not giving the impression that things are going badly amongst other factors [23]. Therefore convenient, is the automatic extraction of information about the relevant variables which permit the attainment of unknown data, which partly determines the efficiency of the stands of a trade fair.

In the Machinery and Tools biennial held in Bilbao in March 2002, information was collected on all these aspects. To be precise, 104 variables of 228 exhibitors were analysed. Of these variables, 7 are continuous and the rest are categorical features, result of an expert discretization. Additionally, for each exhibitor, based on various marketing criteria, the stand's global efficiency was rated as high, medium or low, in terms of the level of achievement of objectives set for the trade fair.

For this real problem, the data mining algorithm should extract information of interest about each of the three efficiency groups of the stands. The rules generated will determine the influence which the different fair planning variables have over the results obtained by the exhibitor, therefore allowing fair planning policies to be improved.

III. Genetic algorithms in rule induction processes

In a data mining process there are different tasks which can be solved as optimisation and search problems. Genetic Algorithms (GAs) are optimisation and search algorithms inspired in natural evolution proceses and initially defined by Holland [15], which have several advantages as a rule induction method (as to cope well with attribute interaction because they usually evaluate a rule as a whole).

In the design of any rule induction GA, the genetic representation of the solutions of the problem is perhaps the most determining aspect of the characteristics of any proposal. The GAs follow two approaches in order to encode rules within a population of individuals:

- The *"Chromosome = Rule"* approach, in which each individual codifies a single rule.
- The *"Chromosome = Set of rules",* also called the Pittsburgh approach, in which each individual represents a rule set. GABIL [6], GIL [17] and GA-MINER [9] are examples of GAs of this type.

In turn, within the "Chromosome = Rule" approach, there are two generic proposals:

- The Michigan approach in which each individual codifies a single rule and a rule set is represented by the entire population. In this case, it is necessary to evaluate the behaviour of the whole set of rules and to define a reinforcement component. The ZCS [26] and XCS [27] algorithms are examples of this approach.
- The IRL (Iterative Rule Learning) approach, in which each chromosome represents a rule, but the GA solution is the best individual obtained and the global solution is formed by the best individuals obtained when the algorithm is run multiple times. In [3] and [12] two proposals with this model are described.
- The "cooperative-competitive" approach, in which the complete population or a subset of it codifies the rule base. COGIN [14], REGAL [11] and [16] are examples of GAs with this type of representation.

IV. Fuzzy logic in rule induction processes

As we know, the principle objective of any data mining process is the identification of interesting patterns and their description in a concise and significant manner. The use of Fuzzy Logic in Data Mining is sensible because fuzzy models represent a description of the data directed towards the user through a set of qualitative models which establish significant and useful relationships between variables. Fuzzy sets allow us to establish flexible limits between the different levels of meaning, without ignoring or overemphasising the elements closest to the edges, as human perception does.

In rule induction processes, Fuzzy Logic is included in such a way that the models extracted are fuzzy rules. In the most interpretable type of fuzzy rules, linguistic fuzzy rules, the continuous variables are defined as linguistic variables. The use of Fuzzy Logic in rule induction processes with quantitative variables eases the interpretability of the knowledge which is finally extracted, the incorporation of qualitative knowledge of the problem, the treatment of lost values and classes with limits which are not well defined, and the processing of noise in variables which are the result of real measurements [1].

One of the fundamental aspects when working with fuzzy rules is the definition of membership functions associated with the fuzzy sets used. In Kouk's algorithm [19] for the extraction of fuzzy rules the expert needs to give the algorithm the continuous variables and their corresponding membership functions. In this case, the quality of the results obtained by the algorithm depends on the suitability of the fuzzy sets. For many applications it is very difficult to know from the outset which fuzzy sets will be the most appropriate. However, in order to increase the interpretability of the results obtained in some proposals such as [1], knowledge of the problem is introduced in the initial definition of the fuzzy sets, such that the rules obtained are based on these fuzzy sets. Our proposal is centred on this approach.

V. An evolutionary approach to obtain descriptive fuzzy rules

In the evolutionary model of extraction of fuzzy rules for subgroup discovery which we present, two components can be distinguished:

- An iterative model of extraction of fuzzy rules for the description of subgroups supported by different areas (not necessarily apart) of the instance space. This model includes the hybrid GA described below.

- A hybrid GA for the extraction of one fuzzy rule that is simple, interpretable, and has an adequate level of support and confidence.

A. Iterative model for extraction of descriptive fuzzy rules

The objective of the model for the extraction of descriptive fuzzy rules is to obtain a set of rules which give information on the majority of available examples for each value of the target variable.

The proposal follows the IRL approach: it includes a hybrid GA which generates a rule in an iterative plan. The iterative model allows new rules to be obtained while the generated rules reach a minimum level of confidence and give information on areas of search space in which examples which are not described by the rules generated by the previous iterations, remain. The algorithm diagram is as follows:

```
START
   RuleSet ← Ø
   REPEAT
    Execute the GA obtaining rule R
    Local Search (R)
    RuleSet ← RuleSet + R
    Modify the set of examples
    WHILE confidence(R) ≥ minimum confidence and
        R represents new examples
END
```

The repetition mechanism promotes the generating of different rules (in the sense that they give information on different groups of examples). This is achieved by penalizing – once a rule is obtained – the set of examples represented by the same in order to generate future rules. It can be consider as a sequential niching GA which remarks differences at phenotypical level.

It is important to point out that the penalization does not impede the extraction of concealed rules. In subgroup discovery algorithms, the possibility of extracting information on described examples is not eliminated since redundant descriptions of subgroups can show the properties of groups from a different perspective. As can be seen in the extraction model diagram, in each iteration the confidence of the obtained rule must be higher than a previously specified minimum value.

B. Genetic algorithm for extraction of a descriptive fuzzy rule

In order to obtain the best fuzzy rule, a hybrid GA which, following the evolutionary obtainment of the fuzzy rule, applies a stage of post-processing, a hill-climbing process is used. The elements of the GA will be described below.

1) Chromosome representation: The objective of the GA is to discover rules whose consequent is formed by a target variable which has been defined previously. The rules generated will be fuzzy or crisp, according to whether the variables involved are continuous or categorical, and are coded according to the "Chromosome = Rule" approach. Only the antecedent is represented in the chromosome and all the individuals in the population are associated with the same value of the target feature. This representation, means that the evolutionary algorithm must be run many times in order to discover the rules of the different classes.

Some of the variables of the problem are continuous variables which are treated as linguistic variables with linguistic labels. The fuzzy sets corresponding to the linguistic labels are defined by a uniform fuzzy partition with triangular membership functions, as shown in Fig. 1.

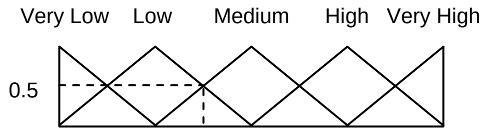

Fig. 1. Example of fuzzy partition for a continuous variable

All the information relating to a rule is contained in a fixed-length chromosome for which we use an integer representation model (the i-th position indicates the value adopted by the i-th variable). The set of possible values for the categorical features is that indicated by the problem plus an additional value which, when it is used in a chromosome, indicates that the corresponding variable does not take part in the rule. For continuous variables the set of values is the set of linguistic terms determined heuristically or with expert information, plus the value indicating the absence of the variable.

Fig. 2 shows an example of the rule and the chromosome which codifies it. In this example, the variable "Employees" does not influence the rule because the possible values are from 1 to 6, and in the corresponding gene the value is 7, which indicates the absence of this variable in the rule.

IF *zone* is *centre* and *sector* is *accessories* and ... and *Bar* is *Yes*
THEN Efficiency is high

Zone	Employees	Sector		Bar
2	7	2	…	1

Fig. 2. Whole encoding model of a rule

2) Fitness function: In this process of rule discovery the objective is to obtain rules with high predictive capacity, and which are understandable and interesting. This objective can be achieved by using a weighted lineal combination of these three measurements [24], as we do in our proposal:

$$fitness\ (c) = \frac{\omega_1 \cdot \text{Support } (c) + \omega_2 \cdot \text{Interest } (c) + \omega_3 \cdot \text{Confidence } (c)}{\omega_1 + \omega_2 + \omega_3}$$

- Confidence: Determines the accuracy of the rule, and reflects the degree to which the examples within the zone of the space marked by the antecedent verify the information indicated in the consequent of the rule. In order to calculate this factor we use an adaptation of Quinlan's accuracy expression [25] to generate fuzzy classification rules [4]: the quotient between the sum of the degree of membership of the examples of this class to the zone determined by the antecedent, and the sum of the degree of membership of all the examples (irrespective of their class) to the same zone. In order to calculate these membership degrees, we use triangular membership functions and the minimum t-norm. In the case of non-fuzzy rules, the degrees of membership correspond to the classic sets, i.e. 0 or 1.

- Support: This is the measurement of the degree of coverage that the rule offers to examples of that class. It is calculated as the quotient between the number of new examples belonging to the class which are covered by the rule and the number of examples (from the same class) which are not covered by the previous extracted rules. This way of measuring support is sensible, when using the GA within an iterative process, in order to obtain different rules each time the GA is run. From the second iteration rules which cover examples belonging to zones delimited by previously obtained rules are penalised, because the support factor only considers examples which have not been described by already-obtained rules. No distance function is used as differences are penalised on a phenotypical level.
- Interest: the degree of interest is assessed objectively. We use the interest criteria provided by [24] in a dependence modelling process, but only using the term referring to the antecedent for the interest calculation, because the consequent is prefixed. The information measurement for the interest is as follows:

$$Interest = 1 - \left(\frac{\sum_{i=1}^{n} Gain(A_i)}{n \cdot \log_2\left(\left|dom(G_k)\right|\right)} \right)$$

where n is the number of variables which appear in the antecendent of the rule, Gain (A_i) is the information gain of the attribute A_i, and |dom(Gk)| is the cardinality (the number of values possible) of the objective variable. Variables with high information gain are suitable for predicting a class when they are considered individually. However, if the user knows the most predictive variables for a specific application domain, the rules containing these variables are less interesting. This way, the antecedent of a rule is more interesting if it contains attributes with a small quantity of information.

The overall objective of the evaluation function is to direct the search towards rules which maximise accuracy, minimising the number of negative and not-covered examples.

3) Reproduction model and genetic operators: A steady-state reproduction model [2] is used: the original population is only modified through the substitution of the worst individuals by individuals resulting from crossover and mutation. The genetic operators used are a multi-point crossover operator and a random mutation operator which is biased such that half the mutations carried out have the effect of eliminating the corresponding variable, in order to increase the generality of the rules.

4) Post-processing phase of the genetic algorithm: local search algorithm: The post-processing phase, which improves the rule obtained by a hill-climbing process, modifies the rule while increasing the degree of support. To accomplish this, in each iteration a variable is determined such that when it is eliminated, the support of the resulting rule is increased; in this way more general rules are obtained. Finally, the optimised rule will substitute the original only if it overcomes minimum confidence. The diagram is as follows:

```
START
Best_ Rule ← R; Best_support ← support(R);
Better ← True
REPEAT WHILE Better
  Better ← False
  FOR (i=1 to gene_number)
    R'i = R without considering variable i
    IF (support (R'i) >= support (R))
      Better ← True
      IF (support (R'i) > Best_support)
        Best_support ← support (R'i)
        Best_Rule ← R'i
  END FOR
  IF (Better AND support(Best_Rule) >= min_supp)
       Return Best_Rule
  ELSE
       Return R
END WHILE
END
```

VI. Experimentation

The experimentation is carried out with the market dataset in which from total set of 104 variables, marketing experts have made a selection of variables which reduces the total set to a subset of 18 variables and the evolutionary rule induction algorithm has been applied to this set of variables.

Parameters of the experimentation:

- The algorithm is run five times for each one of the target variable values.
- Number of chromosomes in the GA: 100.
- Maximum number of evaluations of individuals in each GA run: 5000.
- Fitness function weights. Support: 0.4; confidence: 0.3; interest: 0.3.
- Minimum confidence value: 60.

In Table I the best results obtained are described. Here, for each value of the target variable the confidence, support and interest corresponding to each rule induced are shown (by means of three real numbers belonging to [0,100]). In Tables II, III and IV the rule expressions for efficiency high, low and medium are described.

We can observe that the algorithm induces set of rules with a high confidence (higher than the minimum confidence value) and interest level, around 60 in most cases. This high level of interest, according to the definition of the interest measurement used indicates that the variables which intervene in the general rules are variables with low information gain value, more surprising to the user and they carry more information.

We must note that variables with high information gain are suitable for predicting a class when they are considered individually. However, from the point of view of the interest of a rule, it is understood that the user already knows which are the most predictive variables for a specific application domain, and therefore the rules which contain these variables are less interesting, as they are less surprising and carry less information. Therefore, it is understood that the antecedent of a rule is more interesting if it contains attributes with a small quantity of information, as the rule induces here.

The rule support, except for some rules, is low. The model induces, for this problem, specific rules which represent a small number of examples. The market problem used in this work is a difficult real problem in which inductive algorithms tend to obtain small disjuncts, more common in datasets than one might think at first glance. However, the small disjunct problem, is not a determining factor in the induction process for subgroup discovery. This is because partial relations, i.e., subgroups with interesting characteristics, with a significant deviation from the rest of the dataset, are sufficient.

TABLE I
QUALITY MEASUREMENTS OF THE RULES EXTRACTED

Class	Rule	Support	Confidence	Interest
1	1	10,526	100,000	61,282
	2	13,158	100,000	60,663
	3	18,421	100,000	58,341
	4	7,895	100,000	58,248
	5	7,895	100,000	59,971
	6	5,263	100,000	57,806
	7	5,263	100,000	53,024
2	1	10,811	100,000	59,112
	2	10,135	100,000	55,906
	3	6,081	100,000	58,062
	4	3,378	100,000	61,805
	5	6,081	100,000	59,567
	6	3,378	100,000	57,870
	7	4,730	100,000	59,923
	8	3,378	100,000	60,617
	9	2,027	100,000	60,929
	10	3,378	100,000	59,232
	11	95,946	64,840	62,340
	12	0,676	100,000	60,977
3	1	4,762	100,000	62,110
	2	9,524	100,000	59,904
	3	11,905	100,000	59,045
	4	4,762	100,000	59,845
	5	7,143	100,000	60,580

The knowledge discovered for each one of the target variable values is understandable by the user due to the use of Fuzzy Logic, and the small number of rules and conditions in the rule antecedents.

Marketing experts from Department of Organisation and Marketing of the University of Mondragón (Spain) analysed the results obtained and indicated that:

- The companies which obtain better results (high efficiency) are those that has written objectives, presents authentic innovations in the fair and come from the zone East (Catalonia and Levant). In this regard it must be noted that the exhibitors were coming, principally, from the zone North (where the exhibition was celebrated) and the zone East. Therefore, it can be supposed that the exhibitors proceeding from the zone East, due to the distance, had to do a major economic effort and of time, which pushed to prepare with major attention the participation in the fair.
- On the contrary, the exhibitors who obtained worse results were the manufacturers of the zone North, belonging to the sectors of Deformation and Starting, which had not written objectives and had not done an effort of planning of the campaign of promotion before the event. These exhibitors were not having a list to whom to direct the campaign of promotion before the fair and they did not evaluate the results of the same one, either.

TABLE II
RULES FOR HIGH EFFICIENCY

1	IF Written objectives = Yes AND Stewardesses = No AND Stand at entrance = Yes AND Near of stairs= Yes THEN Efficiency = High
2	IF Sector = Rest AND Number of annual fairs = More than 11 AND New features = Authentic newness THEN Efficiency = High
3	IF Zone = East AND Sector = Rest AND Fairs utility = High AND Importance of contacts quality = High AND New features = Authentic newness THEN Efficiency = High
4	IF Zone = East AND Sector = Rest AND Number of annual fairs = Less than 11 AND Existence of promotion listings = Yes AND Importance of operations after the fair = High AND Quality of contacts= Medium AND Stand at entrance = No THEN Efficiency = High
5	IF Fairs utility = High AND Written objectives = Yes AND New features = Authentic newness AND Stand at entrance = No AND Near of stairs= No THEN Efficiency = High

TABLE III
RULES FOR LOW EFFICIENCY

1	IF Sector = Starting+Deformation AND Written objectives = No AND Previous promotion = No THEN Efficiency = Low
2	IF Written objectives = No AND Importance of present clients contacts = Low AND Quality of contacts= High AND Stand at entrance = No AND Near of stairs= No THEN Efficiency = Low
3	IF Zone = North AND Sector = Starting+Deformation AND Written objectives = No AND Telephone calls = Yes AND New features = Product improvement AND Stand at entrance = No THEN Efficiency = Low
4	IF Importance of contacts = Low AND Quality of contacts= Low THEN Efficiency = Low
5	IF Zone = East AND Written objectives = No AND Existence of promotion listings = No AND Importance of operations after the fair = High AND Stand at entrance = No AND Near of stairs= No THEN Efficiency = Low
6	IF Zone = North AND Fairs utility = Low AND Importance of contacts = Medium AND New features = Product improvement THEN Efficiency = Low
7	IF Sector = Starting+Deformation AND Promotion campaign monitoring = No AND Importance of present clients contacts = High AND Machinery demonstrations type = Sporadic operation AND Stewardesses = Yes THEN Efficiency = Low

TABLE IV
RULES FOR MEDIUM EFFICIENCY

1	IF Zone = North and Fairs utility = Low AND Visitors number importance = Medium AND Stand at entrance = Yes THEN Efficiency = Medium
2	IF Zone = North AND Quality of contacts= High AND Telephone calls = Yes AND New features = "Catalogue" THEN Efficiency = Medium
3	IF Sector = Rest AND Importance of operations after the fair = Medium AND New features = Product improvement THEN Efficiency = Medium
4	IF Sector = Starting+Deformation AND Number of annual fairs = More than 11 THEN Efficiency = Medium
5	IF Previous promotion = Yes AND Visitors number importance = Low AND Stand at entrance = Yes THEN Efficiency = Medium
6	IF Sector = Rest AND Importance of operations after the fair = Low AND Visitors number importance = High THEN Efficiency = Medium
7	IF Zone = North AND Sector = Starting+Deformation AND Fairs utility = Low AND Previous promotion = Yes AND Quality of contacts= Medium THEN Efficiency = Medium
8	IF Quality of contacts= Medium AND Stewardesses = Yes THEN Efficiency = Medium
9	IF Previous promotion = No AND Quality of contacts= High AND Stand at entrance = Yes THEN Efficiency = Medium
10	IF Sector = Rest AND Importance of operations after the fair = Low AND Quality of contacts= Medium THEN Efficiency = Medium
11	IF Number of annual fairs = Less than 11 THEN Efficiency = Medium
12	IF Number of annual fairs = More than 11 AND Quality of contacts= Medium THEN Efficiency = Medium

VII. CONCLUSIONS - CONCLUDING REMARKS

The area of Soft Computing provides a set of tools which, independently or together, are being successfully used in knowledge extraction tasks.

Fuzzy Logic allows the user to incorporate directly linguistic knowledge into the data mining process, to mix this knowledge with non-linguistic information and to treat appropriately incomplete data or data with noise. But perhaps one of the characteristics which is most important for the use of fuzzy logic in this type of algorithm is its ability to represent knowledge in a linguistic form which is directly interpretable, through fuzzy rules.

Genetic Algorithms carry out a global search which is independent of the domain. This makes them a strong tool which can be applied to various stages of the knowledge extraction process.

In this paper we describe an evolutionary model for the descriptive induction of fuzzy or crisp rules which describe subgroups. The proposal includes a GA in an iterative model which extracts rules when some examples are left uncovered, and the rules obtained surpass a given confidence level which is specified by the user. We have applied this proposal to a real knowledge extraction problem in trade fairs. The experiment carried out has determined a simple set of rules which use few variables and therefore has a simple structure. The information extracted is comprehensible for and usable by the final user.

In future studies, we will examine the use of a more flexible structure for the rule and the study of an appropriate interest measurement for this structure. Moreover, we are working in the use of spatial niching GAs and in the development of a multiobjective [7] version of this model to obtain different descriptive rule sets in the pareto-from which can be analysed by the expert.

ACKNOWLEDGMENT

This work was supported by the Spanish Ministry of Science and Technology and by the European Fund. FEDER under Projects TIC-04036-C05-01 and TIC-04036-C05-04.

REFERENCES

[1] W.H. Au, K.C.C. Chan, “An effective algorithm for discovering fuzzy rules in relational databases”, in *Proceedings of IEEE International Conference on Fuzzy Systems (Fuzz IEEE’98)*, pp 1314-1319, 1998

[2] T. Bäck, D. Fogel, Z. Michalewicz, *Handbook of Evolutionary Computation*, Oxford University Press, Oxford, 1997

[3] O. Cordón, F. Herrera, “A three-stage evolutionary process for learning descriptive and approximative fuzzy logic controller knowledge bases from examples”, *International Journal of Approximate Reasoning*, 17 (4), pp 369-407, 1997

[4] O. Cordón, M.J. del Jesus, F. Herrera, “Genetic Learning of Fuzzy Rule-based Classification Systems Co-operating with Fuzzy Reasoning Methods”, *International Journal of Intelligent Systems* 13 (10/11), pp 1025-1053, 1998

[5] O. Cordón, F. Herrera, F. Hoffmann, L. Magdalen, Genetic Fuzzy Systems. Evolutionary tuning and learning of fuzzy knowledge bases, Worldd Scientific, 2001

[6] K.A. De Jong, W.M. Spears, D.F. Gordon, Using genetic algorithms for concept learning, *Machine Learning,* 13 (2-3), pp 161-188, 1993

[7] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*, John Wiley & Sons, 2001

[8] U. Fayyad, G. Piatesky-Shapiro, P. Smyth, P, “From data mining from knowledge discovery in databases”, in U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy (eds), *Advances in Knowledge Discovery & Data Mining,* AAAI/MIT, pp 1-34 , 1996

[9] I.W. Flockhart, N.J. Radcliffe, *GA-MINER: Parallel data mining with hierarchical genetic algorithms* (Final Report by the University of Edimburgh, UK, EPCC-AIKMS-GA-Miner-Report 1.0), 1995

[10] D. Gamberger, N. Lavrac, “Expert guided subgroup discovery: Methodology and application”, *Journal of Artificial Intelligence Research*, 17, pp 501–527, 2002

[11] A. Giordana, F. Neri, “Search-intensive concept induction”, *J Evolutionary Computation*, 3 (4), pp 375-416, 1995

[12] A. González, R. Pérez, “SLAVE: a genetic learning system based on an iterative approach”, *IEEE Trans. Fuzzy Systems* 7(2), pp 176-191, 1999

[13] S. Gopalakrishna, G.L. Lilien, J.D. Williams, I.K. Sequeira, “Do trade shows pay off”, *Journal of Marketing*, 59, pp 75-83 , 1995

[14] D.P. Greene, S.F. Smith, “Competition-based induction of decision models from examples.”, *Machine Learning*, 3, pp 229-257, 1993

[15] J.H. Holland, *Adaptation in natural and artificial systems*. University of Michigan Press, 1975

[16] H. I. Ishibuchi, T. Nakashima, T. Murata, “Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems”, *IEEE Trans. Systems Man Cybernet. B* 29(5), pp 601-618, 1999

[17] C.Z. Janikow, “A knowledge-intensive genetic algorithm for supervised learning”, *Machine Learning*, 13, pp 189-228, 1993

[18] W. Klösgen, “Explora: a multipattern and multistrategy discovery assistant”, in V. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy (eds) *Advances in Knowledge Discovery and Data Mining,* MIT Press, pp 249-271, 1996

[19] C. Kuok, A. Fu, M. Wong, M, "Mining fuzzy association rules in databases", *ACM SIGMOD Record*, 27, pp 41–46, 1998

[20] N. Lavrac, B. Cestnik, D. Gamberger, P. Flach, P, "Decision support through subgroup discovery: three case studies and the lessons learned", *Machine Learning*, 57 (vol 1-2), pp 115-143, 2004

[21] N. Lavrac, P. Flach, B. Kavsek, L. Todorovski, "Adapting classification rule induction to subgroup discovery", in *Proceedings of the Second IEEE International Conference on Data Mining*, IEEE Computer Society, pp 266–273, 2002

[22] M. Mesonero, *Toward an effective trade show planning model based on genetic algorithms*, Ph.D. thesis (in Spanish), Mondragon Unibertsitatea, 2004.

[23] S. Millar, *How to get the most of the trade shows*, NTC Publishing Group, 2003

[24] E. Noda, A.A. Freitas, H.S. Lopes, "Discovering Interesting Prediction Rules with a Genetic Algorithm", in *Proceedings of the Congress on Evolutionary Computation*, 2, pp 1322-1329, 1999

[25] J.R. Quinlan, *Generating production rules. Machine Learning*. Morgan Kaufmann, 1987

[26] S.W. Wilson, D.E. Goldberg, "ZCS: a zeroth order classifier system", *Evolutionary Computation,* 2, pp 1-18, 1994.

[27] S.W. Wilson, "Classifier fitness based on accuracy", *Evolutionary Computation*, 3(2), pp 149-175, 1995.

[28] S. Wrobel S, "An algorithm for multi-relational discovery of subgroups", in *Proceeding of the First European Conference on Principles of Data Mining and Knowledge Discovery*, pp 78-87, 1997