Knowledge Discovery by Genetic Fuzzy Systems Applied to Consumer Behavior Modelling

Jorge CASILLAS
Francisco J. MARTÍNEZ-LÓPEZ
University of Granada, Spain

Abstract

Consumer behaviour discipline has made traditionally use of models to understand consumers. Thus, following the scientific method, marketing academics usually pose theoretical models which are subsequently tested by means of several statistical methods. When such models are complex –i.e. several dependent and independent constructs with multiple relations among them– the method usually used for estimating it is Structural Equation Modelling (SEM). In this sense, the hegemony of SEM for estimating this kind of consumer models has been quite obvious during the last decades. However, we think that this method presents some lacks which constraints its usefulness beyond an academic framework; i.e. it is useful to test models, though results provided by SEM are not good enough for being the kind of support that marketing managers need for guiding their market decisions.

Thus, the main motivation of this paper is caused by a strong belief in the necessity that marketing modelling analytical methods have to evolve, considering the application of other tools of analysis more appropriate to aid the marketing managers' decisional processes.

This paper briefly presents a brand new methodology to be applied in marketing (causal) modeling. Specifically, we apply it to a consumer behavior model used for the experimentation. The characteristics of the problem (with uncertain data and available knowledge from a marketing expert) and the multiobjective optimization we propose make genetic fuzzy systems a good tool for tackling it. In sum, by applying this methodology we obtain useful information patterns (fuzzy rules) which help to better understand the relations among the elements of the marketing system (causal model) being analyzed; in our case, a consumer model.

Keywords: marketing modeling, decisions support systems, knowledge discovery methodology, genetic fuzzy systems, consumer's behavior patterns.

1. Background

Marketing academics and practitioners have pointed out the necessity for knowing and explaining the consumer's behaviour patters in a manner increasingly efficient. Firms focused on final markets are immersed in highly competitive systems in which it is needed that their decision processes to be as correct as possible.

In this regard, models of consumer behaviour, inasmuch as they are marketing models, are considered as a specific case of Marketing Management Support System (MkMSS), and throughout the time have demonstrated to be a source of transcendental relevance for the development of marketing science (van Bruggen & Wierenga, 2000).

Notwithstanding, current models of consumer behaviour do not seem to cover all the necessities that it should supposedly satisfy a model which aims to aid on the marketing decision making. With respect to this, based on Gatignon (2000), future models, considering both their theoretical and technical aspects, which try to explain consumers' decision making will have to be clearly focused on users' (demand side) requirements of such models. That is to say, models must be more complete, flexible, and customized to the strategic singularities of the competitive environment which their users operate in. Thus, as the main problem that actually face firms oriented to consumer markets is not the availability of information (data), but the possession of necessary level of knowledge to take the right decisions, the use of avant-garde behavioural models able to exploit it may represent an essential source of competitive advantage.

Doubtless, it is time to rethink the role of marketing research and modelling, in order to develop more adequate analytical methods to tackle the current business environments and decisional scenarios (Wind, 2006). It is expected that MkMSS will tend to improve their performance taking advantage of synergies caused by the integration of modelling estimation techniques based on classic econometric with expert systems based on artificial intelligence.

Specifically, considering the three pillars in which marketing modelling is based (Roberts, 2000), and more specifically the consumer behaviour modelling, we focus our paper on one of them, i.e.: the modelling estimation techniques and its improvement. We treat the potentials that analytical methods based on fuzzy rules have to evolve the method of estimation and analysis traditionally used till now, basically based on classic statistical (parametric) techniques. In this sense, fuzzy rules can be a plausible alternative or complement to the results obtained by using Structural Equation Modelling (SEM) techniques which have been the ones usually used in the last decades to estimate complex models of consumer behaviour.

This paper presents a Knowledge Discovery in Database (KDD) methodology developed *ad hoc* to be applied in marketing (causal) modeling. A *descriptive rule* *induction* method (Lavrac et al., 2004) is posed to discover individual rules which show information patterns of especial interest in the data. To do this, we consider fuzzy association rules, but previously setting antecedents' and consequents' variables; i.e. we use a theoretic (causal) model of reference, which is used to supervise the machine learning process. Extraction is realized by genetic fuzzy systems, a soft computing hybridization. An empirical illustration of how it works is also provided.

2. Knowledge Discovery Based on Fuzzy Rules

In general terms, knowledge discovery in databases (KDD) is a recent research field belonging to artificial intelligence whose main aim is the identification of new, potentially useful, and understandable patters in data (Fayyad, Piatesky-Shapiro, Smyth & Uthurusamy, 1996). Furthermore, KDD implies the development of a process which is compounded by several stages. In this sense, data mining, which is considered as the core of KDD process, is characterized by the application of machine learning methods to automatically or semi-automatically extract patterns or models from data (Witten & Frank, 1999).

Nowadays, one of the most successful tools to develop descriptive models is fuzzy modelling (Lindskog, 1997), which is an approach used to model a system making use of a descriptive language based on fuzzy logic with fuzzy predicates (Sugeno & Yasukawa, 1993). The way to express fuzzy predicates is by means of IF-THEN rules with the following structure:

IF X_1 is A_1 and … and X_n is A_n THEN Y_1 is B_1 and … Y_m is B_m

These rules set logical relationships among variables of a system by using qualitative values. Such representation mode has the power to be easily understandable by human being. Hence, the performance of both, analysis and interpretation steps of the modelling process, improve thanks to the true behaviour of system is more effectively revealed. Notwithstanding, it should be noted that though human reasoning may manage without strain with terms like *high* or *will rise quickly*, when this issue is tackled by means of an automatic process its treatment is more complex.

To properly work with this kind of qualitative valuations, linguistic variables (Zadeh, 1975) based on both Fuzzy Sets Theory and Fuzzy Logic (Zadeh, 1965) are used, so previous exemplified rule is known as a *fuzzy rule*. The use of fuzzy logic provides several benefits as: a higher generality, expressive power, ability to model real problems and, at last but not least, a methodology to exploit tolerance in the face of imprecision. In example, we can consider the linguistic variable *age*, which linguistic values could be *teenager*, *young*, *adult*, and *old*.

Fuzzy rules can be considered as a knowledge extraction tool to discover intrinsic relationships contained in a database (Freitas, 2002). Thus, by means of fuzzy rules we can represent the relationship existing among different variables, thus deducing the patterns contained in the examined data. In knowledge discovery, the process to obtain these patterns must be automatic, or semi-automatic, discovered patterns must be comprehensible and they must provide useful information, and data must be invariably presented in substantial quantities (Witten & Frank, 2000).

Useful patterns allow us to do non trivial predictions about new data. There are two extremes to express a pattern: like black boxes, whose internal behaviour is incomprehensible; and like white boxes, whose construction reveals the pattern structure. The difference lies in whether the generated patterns are represented with an easily examined structure, which can be used to reason and to inform further decisions. In other words, when the patterns are structured in a comprehensible way, they will be able to help in explaining something about the data. This trouble of KDD, the interpretability-accuracy trade-off, is also being currently faced in fuzzy modelling (Casillas *et al.*, 2003a, 2003b) and will be considered by our proposal.

The use of fuzzy rules when developing the knowledge discovery process has some advantages as follows: they allow us to use uncertainty data; they adequately consider multi-variable relationships; results are easily understandable by a human being; additional information can be easily added by an expert; the accuracy degrees can be easily adapted to the problem necessity; and the process can be highly automatic with low human intervention.

Therefore, we will use fuzzy logic as a tool to structure the information of a consumer behaviour model in a clear, legible, and close to the human being way. The fuzzy system will allow us to properly represent the interdependence of variables and the non-linear relationships that could exist among them. Finally, optimization algorithms (a genetic algorithm in this paper) will design the fuzzy rules to meet the interpretability and accuracy criteria imposed by the expert.

The following section introduces the methodology followed for applying data mining by means of fuzzy rules to consumer behaviour modelling.

3. A Methodology for Consumer Behavior Modelling by Genetic Fuzzy Systems

3.1. Data Gathering

First step is to collect the data related to the variables defining the theoretic consumer behavior model of reference. In this sense, as it has been traditionally done in marketing, data are obtained by means of a questionnaire. Thus, firstly, attention should be paid to how consumer behavior modelers face and develop the measurement process of variables that complex behavioral models contain; i.e. usually, latent/unobserved variables. Its understanding is necessary in order to adequately approach the starting point of the KDD process, so to give suitable and adapted solutions to the specific data we find in consumer behavior modeling.

It can be said that measuring streams for these latent variables in marketing modeling can be classified into two groups depending on if they state that these constructs can or cannot be perfectly measured by means of observed variables (indicators); i.e., the existence or not of a one-to-one correspondence between a construct and its measurement. Certainly, though consumer behavior modelers tended to make use in the beginning of what was known as the *operational definition philosophy*, a more convenient and reasonable position is that ulteriorly based on the *partial interpretation philosophy* which distinguished between unobserved (constructs) and observed (indicators) variables. This latter approach of measurement, being currently predominant in the marketing modeling discipline, poses to jointly consider multiple indicators –imperfect when considered individually, though reliable when considered altogether– of the subjacent construct to obtain valid measures (Steenkamp & Baumgartner, 2000). Hence, we will take this measurement approach into account when facing how to process the data.

3.2 Data Processing

Next, it is necessary to adapt the collected data to a scheme easily tractable by fuzzy rule learning methods. Therefore, our methodological approach should be aware of the special features of the available data (with several items or indicators to describe a specific variable) when adapting the observed variables to a fuzzy rule learning method. An intuitive approach could directly reduce the items of certain variables to a single value (e.g., by arithmetic mean). Another possibility would be to expand any multi-item example (the result of a questionnaire filled out by a consumer) to several single-item examples and, subsequently, reduce the data size with some instance of selection process (Casillas, Martínez-López & Martínez, 2004).

The problem of these approaches is that the data must be transformed, so relevant information may be lost. We propose a more sophisticated process that allows working with the original format without any pre-processing stage: the *multi-item fuzzification*. Thus, a *T-conorm* operator (e.g., maximum), traditionally used in fuzzy logic to develop the union of fuzzy sets, is applied to aggregate the partial information given by each item during the inference process. Since it is not pre-processing data but a component of the machine learning design, the details of that treatment of the items is described in Section 3.4.2.

3.3. Representation and Inclusion of Expert Knowledge

Several issues should be tackled at this step: the set of variables to be modeled, the transformation of marketing scales used for measuring such variables into fuzzy semantic and the fuzzy rule structure (relations among constructs). We suggest some approaches to fix these components. All of them are based on the marketing expert's capability to express his knowledge in a humanly understandable format by fuzzy logic.

3.3.1. Fuzzy Semantics from Expert Knowledge

Once the marketing modeler has finally determined both, the theoretical constructs and the observed variables associated with each one (i.e. the measurement model), a transformation of the original marketing scales used for measuring those observed variables into linguistic terms should be done. At this point, several marketing scale types can be used for its measurement. With the aim of simplifying the problem, in this paper we focus on Likert-type, differential semantic and rating scales, which are the most commonly used in these models. The transformation should be practiced taking into account three main questions:

1. The *number of linguistic terms* to be used for each variable must be defined. An odd number seems to be a good approach since in our case it is useful to linguistically express the "medium" or "unconcerned" concept. Since traditional interval scales used in marketing usually present between 5 to 9 different degrees (i.e. points of the scale), the use of three or five linguistic terms (fuzzy sets) is enough to map these values.

2. The *membership function type* defining the behavior of certain fuzzy variables should be also defined. In this sense, such behavior can be broadly treated considering the use of linear (trapezoidal or triangular) vs. non linear (Gaussian) membership functions to characterize the fuzzy sets. In this respect, we pose that it is more appropriate to use linear functions, inasmuch as it facilitates the latter interpretation of relations.

3. The *membership function shapes* should also be fixed. In this respect, we propose to impose some properties in order to ensure good interpretability. Extreme values of the interval should have a membership degree 1 to extreme labels. Mean value of the interval should have membership 1 to medium label. Likewise, we consider strong Ruspini's fuzzy semantics (the sum of the membership degrees of every value to the set of linguistic terms is 1) in order to ensure good interpretability. Finally, in order to statistically unbias the significance of every linguistic term, we impose the same covering degree. Thus, we define the membership function shapes where, given the set S={min,…,max} defining the interval, they hold the following condition:

$$\sum_{k \in S} \mu_{A_i}(k) = \frac{\max - \min}{l}, \ \forall A_i \in A,$$

with l being the number of linguistic terms and $A = \{A_1, \ldots A_l\}$ the set of them.

To sum up, Figure 1 shows an example based on the transformation of a nine-point rating scale (a typical marketing scale used to measure the observed variables/indicators related to certain construct) into a fuzzy semantic with the three linguistic terms *Low*, *Medium*, and *High*.

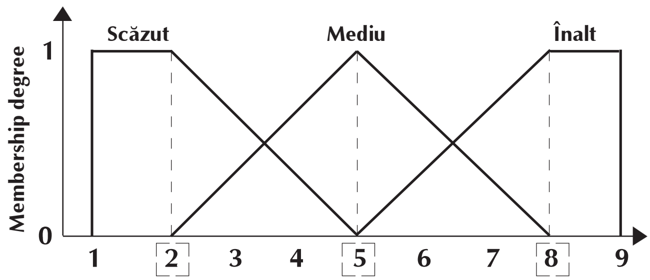

Fig. 1. Fuzzy semantic from a transformation of a 9-point marketing scale (rating scale)

3.3.2. Input/Output Linguistic Variables from Expert Knowledge

Furthermore, once the structure of the model has been fixed by the marketing expert under the base of the theoretic model, fuzzy rules are used to relate input (antecedents) with output (consequents) variables. Obviously, hypotheses contained in the model can be directly used to define IF-THEN structures by considering the dependencies shown among the variables. Thus, we obtain a fuzzy rule base for each consequent (endogenous construct) considered and its respective set of antecedents.

For example, if we took for illustrative purposes the model associated with the Theory of Reasoned Action (Ajzen & Fishbein, 1980), the fuzzy rule structure which represents the widely known relations between the elements "attitude" and "subjective norm" with the consequent "intention" will have the following form:

IF *Attitude* is A_1 and *SubjectiveNorm* is A_2 **THEN** *Intention* is B.

3.4. Data Mining Process

Once the linguistic variables that properly represent the tackled information have been fixed, a machine learning process must be used to automatically extract the knowledge existing in the database. This process is, without any doubt, the most important issue from the KDD point of view.

As mentioned in the Background Section, in this paper we are interested in descriptive induction. Therefore, we will use GAs Michigan-style to obtain rules individually relevant. We consider two quality criteria, support (degree of representativity of the rule with respect to the set of data) and confidence (degree of accuracy of the relation shown by the rule). It is intuitive to check that the higher the support, the higher the difficulty to maintain high degrees of confidence. To jointly consider both criteria, we propose the use of *multiobjective GAs*, as they offer good results when working with multiple contradictory objectives. The next section describes the main elements of this method we propose.

3.4.1. Fuzzy Rule Structure

In data mining it is crucial to use a learning process with a high degree of interpretability. To do that, we opt for a compact description based on the disjunctive normal form (DNF). This kind of fuzzy rule structure has the following form:

IF X_1 is $\widetilde{A}_1$ and … and X_n is $\widetilde{A}_n$ **THEN** Y_1 is B

where each input variable X_i, $i \in \{1,\ldots,n\}$ takes as a value a set of linguistic terms $\widetilde{A}_i = \{A_{i1}$ sau … sau $A_{in_i}\}$, whose members are joined by a disjunctive operator. We use the bounded sum $\min\{1,a+b\}$ as *T-conorm*. The structure is a natural support to allow the absence of some input variables in each rule, simply making $\widetilde{A}_i$ to be the whole set of linguistic terms available.

3.4.2. Multi-item Fuzzification

In order to properly consider the set of indicators available for each input/output variable (as discussed in Section 3.2), we propose an extension of the membership degree computation, the so-called *multi-item fuzzification*. The process is based on a union of the partial information provided by each item. Given X_i and Y_j measured by the vectors of items $\vec{x}_i = (x_1^{(i)},\ldots, x_{h_i}^{(i)},\ldots, x_{p_i}^{(i)})$ and $\vec{y} = (y_1,\ldots, y_t,\ldots, y_q)$, respectively, the fuzzy propositions X_i is $\widetilde{A}_i$ and Y is B are respectively interpreted as follows:

$$\mu_{\tilde{A}_i}(\vec{x}_i) = \min\left\{1, \bigcup_{h_i=1}^{p_i} \sum_{A\in\tilde{A}_i} \mu_A(x_{h_i}^{(i)})\right\}$$

$$\mu_B(\vec{y}) = \bigcup_{t=1}^{q} \mu_B(y_t),$$

with $\bigcup$ being a T-conorm (the maximum in this paper).

3.4.3. Subgroup Discovery

To do the descriptive rules induction process, we have applied a method with certain similarities to the subgroups discovery technique –widely used in classification learning rules (Lavrac, 2004)–, where the property of interest is the class associated with the variables of the consequent. Therefore, we try to group the set of data into differentiated subgroups, including in each of them those examples represented by the consequent with the aim of discovering a representative set of rules for each subgroup. In this regard, the most usual approach is based on running the algorithm designed for each subgroup of data which satisfies the property set for the consequent.

However, instead of this approach, we carry out a simultaneous subgroup discovery in the algorithm we propose. This variant allows us to form niches of fuzzy rules differentiated by the consequent which are optimized in parallel to finally generate a set of suboptimal solutions for each class of the consequent. With the aim of developing this simultaneous process, as it is shown in the next sections, we vary the concept of multiobjective dominance by making the genetic operators act only on the antecedents of the rules.

3.4.4. Coding Scheme

Each individual of the population represents a fuzzy rule; i.e. a Michigan-style genetic algorithm. The coding scheme will be binary to represent the antecedent and whole for the consequent. Thus, the allele “1” in the antecedent part means that the linguistic term related to the gene is used in the corresponding variable. For the consequent, we will directly code the index of the linguistic term used. Hence, the size to code a DNF fuzzy rule is equal to the sum of the number of linguistic terms employed in each input variable (antecedent) plus the number of output variables. For instance, if we had three linguistic terms for each variable, the rule [IF X1 is Small and X2 is {Medium or High} THEN Y is Medium], would be coded as [100 011|2].

3.4.5. Objective Functions

In this algorithm, we consider the two criteria most frequently used to value the quality of the association rules (Dubois, Prade & Sudkamp, 2005): support and confidence. However, we adapt the calculus of these criteria to fuzzy association rules, also considering the especial characteristics of the multi-item variables (elements of the model) which we work with.

Support. This objective function values the degree of representation of certain fuzzy rule on the set of data analyzed. It is calculated as the average degree covered by the rule considering every one of these data (individuals´ responses). To obtain the degree of cover we conjointly consider the membership degrees in relation to the diverse variables; i.e. the set of antecedents as well as the consequent. The measure of support (for maximization) for a fuzzy rule R comes defined as follows:

$$Support(R) = \frac{1}{N}\sum_{e=1}^{N} T\,(\mu_A(x^{(e)}),\ \mu_B(\vec{y}^{(e)})),$$

where N is the size of the database (the sample size or number of respondents), $x^{(e)} = (\vec{x}_1^{(e)},\ldots,\vec{x}_n^{(e)})$ and $\vec{y}^{(e)}$ is the eth instance multi-item of input and output respectively, T the *product* T-norm, and $\mu_A(x^{(e)}) = \min_{i\in\{1,\ldots,n\}} \mu_{\tilde{A}_i}(\vec{x}_i^{(e)})$ the coverage degree of the antecedent of the rule R for this example (i.e. it is considered the T-norm of the minimum to interpret the connector "and" of the fuzzy rule). Also, it is convenient to point out that we employ the multi-item fuzzification shown in section 3.4.2 to calculate $\mu_{\tilde{A}_i}(\vec{x}_i^{(e)})$ and $\mu_B(\vec{y}^{(e)})$.

Confidence. This objective function measures the reliability of the relationship between antecedent and consequent described by the analyzed fuzzy rule. We have used a confidence degree that avoids accumulation of low cardinalities [4]. It is computed (for maximizing) as follows:

$$Confidence(R) = \frac{\sum_{e=1}^{N} T\,(\mu_A(x^{(e)}), I(\mu_A(x^{(e)}), \mu_B(\vec{y}^{(e)})))}{\sum_{e=1}^{N} \mu_A(x^{(e)})},$$

The Dienes' S-implication $I\,(a,b) = \max\{1 - a, b\}$ is used. We consider again T-norm of product and multi-fuzzification.

3.4.6. Evolutionary Scheme

A generational approach with the multi-objective NSGA-II replacement strategy (Deb et al., 2002) is adopted. A binary tournament selection is used based on the crowding distance in the objective function space. To correctly develop the simultaneous subgroup discovery we will need to redefine the concept of dominance. In order to do this, one solution (rule) will dominate another when, as well as equaling as minimum all the objectives and improving in one of them, it presents the same consequent as the other rule. Hence, those rules with different a consequent do not dominate each other. Consequently, we force the algorithm to form so many niches of search (Pareto sets) as diverse consequents (subgroups) are considered.

3.4.7. Genetic operators

The initial population is built defining so many groups (equal in size) as there are different consequents. In each of them, chromosomes are generated fixing such consequents and randomly building a simple antecedent where each input variable is related to a linguistic term. The two operators of reproduction only act in the part of the antecedent of the rule. This fact ensures that the size of every subgroup in the population is constant. In this way, we allow the algorithm to independently explore, but simultaneously, each group.

We employ a multipoint crossover operator which selects two crossover points (in the part of the antecedent) and interchanges the central sub-chain. The operator of mutation randomly selects a variable of the antecedent of the fuzzy rule coded in the chromosome and carries out some of the three following operations: *expansion*, which flips to 1 a gene of the selected variable; *contraction*, which flips to 0 a gene of the selected variable; or *shift*, which flips to 0 a gene of the variable and flips to 1 the gene immediately before or after it. The selection of one of these mechanisms is made randomly among the available choices (e.g., contraction cannot be applied if only a gene of the selected variable has the allele 1)

4. Empirical illustration of the methodology's performance

4.1 Previous commentaries about the model used for the empirical illustration

The experimentation of the descriptive rule induction method we present has been made based on a causal model already proposed by Novak, Hoffman & Yung (2000). It analyzes the consumer's flow state in interactive computer-mediated environments.

In order to briefly introduce this concept, so the reader better understands the variable we want to explain in this empirical application of our methodology, we now synthetically present some ideas about it. Flow has been recently imported from motivational psychology and successfully adapted to explain consumer behavior phenomena on the Web (Hoffman & Novak, 1996; Korzan, 2003; Luna, Peracchio & De Juan, 2002; Novak, Hoffman & Duhachek, 2003; Novak, Hoffman & Yung, 2002).

In general terms, flow state is defined as “the process of optimal experience” or the mental state that individuals sometimes experience when they are deeply immersed in certain events, objects or activities (Csikszentmihalyi, 1975, 1977). This concept has been adapted to the Web environment. In this context, flow state is achieved when the consumer is so deeply involved in the process of navigation on the Web that “nothing else seems to matter” (Hoffman & Novak, 1996: p. 57)

Though the model we consider for the experimentation has 12 elements (constructs) interconnected, with 6 fuzzy rule based systems, due to the space constraints, in this paper we focus on that system which considers the four primary antecedents of the consumer’s flow. Specifically, we consider four constructs (speed of interaction, skill/control, challenge/arousal and telepresence/time distortion) as antecedents of the consumer´s flow state (consequent). In this sense, it is been hypothesized that these four elements are positively related to this central construct of the model.

Most parts of the construct, except one of them which was measured by means of an ordinal scale, were gathered by multi-item Likert scales with 9 points; i.e. metric scales. The fuzzy semantic we have applied to all the variables is shown in figure 1.

Training data are composed of 1,154 examples (consumers´ responses). We have run the algorithm 10 times, obtaining the following values for the parameters: 300 generations, size of the population 100, crossover probability 0.7 and the probability of mutation per chromosome 0.1.

4.2. Analysis of the Pareto Front

The Pareto front we have obtained is shown in Figure 2. With respect to the value taken by the consequent flow in the rules generated, it can be easily observed that the most plausible output is “medium”. Indeed, there is a clear supremacy of the rules with this label in the consequent over the two other outputs in terms of support and confidence. This fact is intensified as the support of the rules grows, without noticing a relevant loss of reliability in the rules which represent medium flow states. Therefore, it can be inferred that the most representative state of flow, for the whole consumers´ database, is moderate.

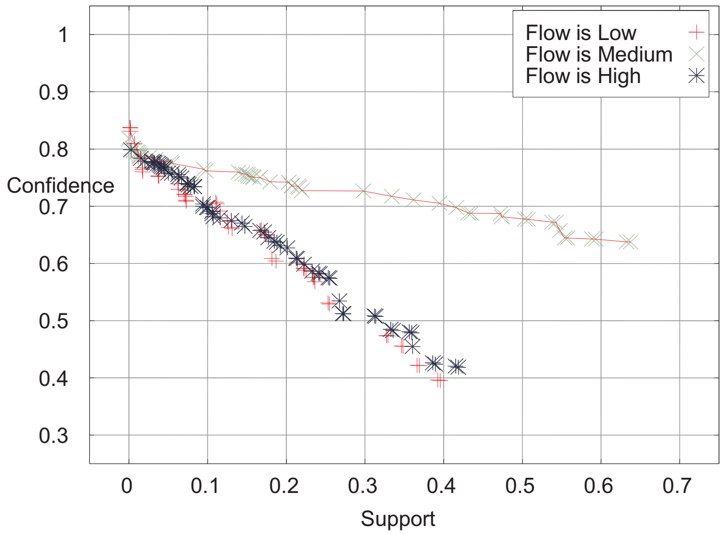

Fig. 2. Sub-Pareto fronts for every output of the consequent, as well as the absolute Pareto front (the best rules from the whole set of rules) joined by a line.

4.3. Illustrative Analysis of the Rules

An individual analysis of the rules generated by this descriptive method is very useful to better understand the consumer behavior being analyzed. Specifically, it is recommendable to do a selection of rules from the whole set compounding the absolute Pareto front, paying attention to its support (degree of representativity of the consumers´ database) and, especially, to its confidence (degree of reliability of the information pattern shown by the rule). In this regard, we have done an illustrative selection shown in Table 1.

Table 1. Illustrative selection of rules from the absolute Pareto front

	Speed of Interaction	*Skill/Control*	*Challenge/ Arousal*	*Telepresence/ Time Distortion*	*Flow*	*Sup*	*Conf*
R_1	Low High	Medium		Low	Low	0,0104	0,7980
R_2	Medium	Low High	High	Medium	Medium	0,0102	0,7937
R_3	Medium			Medium High	Medium	0,3947	0,7051

Considering the absolute Pareto front, R_1 is the rule with highest confidence, associated with low states of flow. Likewise, R_2 represents the most reliable rule from those with moderate flow states. Finally, we have also considered the rule R_3, being the one with highest support among the whole set of rules with confidence higher than 0,7; i.e. the confidence threshold value we have set to give reliability to the information patterns shown by the rules.

Synthetically, from the four antecedents considered, it highlights the influence of the perception about telepresence/time distortion (TP/TD) in determining consumers´ states of flow; it can be observed how its value is determinant in explaining low (R_1) or moderate (R_2 and R_3) states of flow. Likewise, the rest of the antecedents seem to exert a poor or null influence on the consequent. This fact can also be due to the element TP/TD that eclipses the influence of the rest. In any case, it conforms to the main idea we extracted when the Pareto front was analyzed; i.e. a non existence of combinations of antecedents (rules) producing high states of flow, with significant levels of reliability and representativity. In this sense, it is quite illustrative to see how even when the most influential antecedent (TP/TD) takes high values, the consumer's flow state in the process of navigation tends to remain moderate.

5. Final Remarks

An analytical method for estimating complex consumer behavior models should not be only useful to test a set of theoretical relations compounding such model. Moreover, it must be also able to be helpful for the marketing management function to have a good perspective of certain consumption situation, so to take the right decisions. Marketing researchers, especially those focused on improving and developing the "arsenal" of the marketing modeling tools, must be aware of this, in order to bring the gap, with their proposals, between the academics' and the professionals' arenas.

We have presented a complete methodology to be applied in causal marketing modelling by a genetic fuzzy system, a specific soft computing hybridization, with a fuzzy rule descriptive induction approach. This method allows the researcher to obtain a view of the relations among variables in a new way, when compared with the kind of output we use to obtain relations from the statistical techniques in our discipline. It offers singular information patterns for every causal relation contained in the theoretical model used to guide the machine learning process. In this regard, such a process is driven by a genetic algorithm with a multiobjective optimization approach, especially designed for proper management with the kind of measurement scales used in marketing. Furthermore, due to the benefits provided by fuzzy logic, such patterns are expressed in an easily understandable way regarding the way human beings reason.

Acknowledgments

This research has been supported in part by the Spanish Ministry of Education and Science under research project TIN2005-08386-C05-01. A shorter version of this paper was presented at the 2007 EMAC Conference.

References

Ajzen, I., & Fishbein, M.(1980). Understanding Attitudes and Predicting Social Behavior. Englewood Cliffs, NJ: Prentice-Hall,Inc.

Casillas, J., Cordón, O., Herrera, F., Magdalena, L. (Eds) (2003a). *Interpretability issues in fuzzy modeling*. Springer, Heidelberg, Germany.

Casillas, J., Cordón, O., Herrera, F., Magdalena, L. (Eds) (2003b). *Accuracy improvements in linguistic fuzzy modeling*. Springer, Heidelberg, Germany.

Casillas, J., Martínez-López, F.J., & Martínez, F.J. (2004). Fuzzy association rules for estimating consumer behaviour models and their application to explaining trust in Internet shopping. *Fuzzy Economic Review*, IX(2), 3-26.

Csikszentmihalyi, M. (1975). Play and intrinsic rewards. *Journal of Humanistic Psychology*, 15(3), 41-63.

Csikszentmihalyi, M. (1977). Beyond boredom and anxiety. Second edition. San Francisco: Jossey-Bass.

Deb, K., Pratap, A., Agarwal, S., & Meyarevian, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Computation*, 6(2), 182-197.

Dubois, D., Prade, H., Sudkamp, T. (2005). On the representation, measurement, and discovery of fuzzy associations. *IEEE Trans. Fuzzy Systems*, 13(2), 250–262.

Fayyad, U.M., Piatesky-Shapiro, G., Smyth, S., Uthurusamy, R. (1996). *Advances in knowledge discovery and data mining*. M.I.T. Press.

Freitas, A.A. (2002). *Data mining and knowledge discovery with evolutionary algorithms*. Springer, Heidelberg, Germany.

Gatignon, H. (2000). Commentary on Peter Leeflang and Dick Wittink's "Building models form marketing decisions: past, present and future". *International Journal of Research in Marketing*, 17, 209–214.

Hoffman, D., Novak, T. (1996). Marketing in hypermedia computer-mediated environments: conceptual foundations. *Journal of Marketing* 60(July), 50-68.

Korzaan, M.L. (2003). Going with the flow: predicting online purchase intentions. *Journal of Computer Information Systems*, Summer, 25-31.

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

Lindskog, P. (1997). Fuzzy identification from a grey box modeling point of view. In Hellendoorn & Driankov (Eds.), *Fuzzy model identification*, pp. 3-50. Springer-Verlag, Heidelberg, Germany.

Luna, D., Peracchio, L.A., De Juan, M.D. (2002). Cross-cultural and cognitive aspects of Web site navigation. *Journal of the Academy of Marketing Science,* 30(4), 397-410.

Novak, T., Hoffman, D., Duhachek, A. (2003). The influence of goal-directed and experiential activities on online flow experiences. *Journal of Consumer Psychology*, 13 (1/2), 3-16.

Novak, T., Hoffman, D., Yung, Y. (2000). Measuring the customer experience in online environments: A structural modeling approach. *Marketing Science*, 19(1), 22-42.

Roberts, J.H. (2000). The intersection modelling potential and practice. *International Journal of Research in Marketing* 17, pp. 127-134.

Steenkamp, J., Baumgartner, H. (2000). On the use of structural equation models for marketing modeling. *International Journal of Research in Marketing*, 17, 195–202.

Sugeno, M., Yasukawa, T. (1993). A fuzzy-logic-based approach to qualitative modeling, *IEEE Transactions on Fuzzy Systems* 1(1), pp. 7-31.

Van Bruggen, G.H., Wierenga, B. (2000), "Broadening the perspective on marketing decision models", *International Journal of Research in Marketing*, vol. 17, pp. 159-168.

Wind, J. (2006). Challenging the mental models of marketing. In J.N. Sheth and R.S. Sisodia (Eds.) *Does marketing need reform? Fresh perspectives on the future*, M.E. Sharpe, pp. 91-104.

Witten, I.H., Frank, E. (1999), *Tools for data mining: practical machine learning tools and techniques with Java implementations*, Morgan Kauffman Publishers, San Francisco, CA, USA..

Witten, I.H., Frank, E. (2000), *Data mining: practical machine learning tools and techniques with Java implementations*, Morgan Kaufmann Publishers, San Francisco, CA, USA.

Zadeh, L.A. (1965). Fuzzy sets, *Information and Control* 8, pp. 338-353.

Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning. Parts I, II and III, *Information Science* 8, 8, 9, pp. 199-249, 301-357, 43-80.