CS-BIGS 4(2): 108-115

http://legacy.bentley.edu/csbigs/documents/Guillen.pdf

Determining Marital Dissolutions Duration with Fuzzy Inference Systems

Alberto Guillen
University of Granada, Spain

Federico Montesino
Aalto University, Finland

Carlos Tovar
University of Granada, Spain

Luis-Javier Herrera
University of Granada, Spain

Hector Pomares
University of Granada, Spain

Angel Barriga
University of Seville, Spain

Jose Guillen
University of Granada, Spain

Ignacio Rojas
University of Granada, Spain

This paper tackles the problem of predicting how long a martial dissolution process will last. As previous studies in the sociological and sanitary areas show, these types of separations can damage the health and economy of the people involved in the process. Furthermore, the administration overload due to these dissolutions can make the separation process last even longer, causing a deeper trauma. Therefore, it is useful to have a model that is able to predict the duration in order to prevent mental illness and to assign administration resources to speed up the paperwork. The model employed in this paper after performing a dimensionality reduction is a Fuzzy Inference System with a fuzzy-possibilistic algorithm to obtain the values of its parameters.

1. Introduction

Marital dissolutions have many undesirable consequences for society, from the point of view of public health and given the increase in resources that have to be assigned to deal with them. With the introduction of divorce in Spain on June 22[ed] 1981, it has become possible to start measuring and to study consequences for the people involved. Preliminary studies conclude that a dissolution has several impacts on the quality of life and welfare of those involved (Holmes & Rahe, 1967; Sandín & Chorot, 1996; Organización Panamericana de la Salud., 2003; World Health Organization., 2005). These impacts have been analyzed by the European Union in its latest report European Commission (2009) where it was shown that the probability of suffering a mental illness was higher for divorced or separated people than for married ones. In this report, Spain, with an odds ratio of 2.9 was the second country in descending order. These results were corroborated by another study on mental problems in Spain (Haro et. al., 2006).

Health problems are not only faced by the couple, if there are children involved in the process, they can experience low levels of adaptation and integration into society and less healthy life habits in comparison with children who have grown up in regular families (Fariña et. al., 2003; Orgilés et. al., 2008; Marí-Klose et. al., 2009). In order to avoid these problems, as suggested by Lansford (2009), children should not be implied directly in the process and should have a continuous and fluent relationship with their families.

It is quite useful to be able to predict how long the dissolution process might last in order to give some psychological assistance to the children and their parents and to assign more resources to potentially long duration dissolutions in order to accelerate them. In order to do so, this paper proposes a classical methodology to build a regression model: perform a dimensionality reduction and afterwards design the model. The main novelty is the algorithm used to design the model, basing its approach on a hybrid fuzzy-possibilistic partition of the data to decrease the influence of outliers in the data set.

2. Methodology

This section describes the method used to determine which of the available variables are more relevant to build an accurate model. The model selected is then described as well as the algorithm which was used in the training stage.

2.1. Dimensionality Reduction

In order to reduce the dimensionality, Mutual Information (MI) theory is used. The mutual information (also called cross-entropy) between X and Y can be defined as the amount of information that the group of variables X provide about Y, and can be expressed as $I(X,Y) = H(Y) - H(Y \mid X)$ where H denotes the entropy. In other words, the mutual information I(X, Y) is the decrease in the uncertainty on Y once we know X. Due to the properties of mutual information and entropy, the mutual information can also be defined as $I(X,Y) = H(X) + H(Y) - H(X \mid Y)$ leading to:

$$I(X,Y) = \int \mu_{X,Y}(x,y) \log \frac{\mu_{X,Y}(x,y)}{\mu_X(x)\mu_Y(y)} dxdy \quad (1)$$

where $\mu_{X,Y}(x,y)$ is the joint Probability Density Function (pdf) for (X,Y). Thus, only the estimate of the joint Probability Density Function (PDF) of (X, Y) is needed to estimate the mutual information between two groups of variables.

Estimating the joint probability distribution can be performed using a number of techniques such as histograms and kernel density estimators. This paper uses the method based on k-nearest neighbors presented by Kraskov et. al. (2004). As recommended by Harald et. al. (2004) for a trade-off between variance and bias, in the examples, a mid-range value for $(k = 6)$ will be used.

2.2. Fuzzy Inference System Identification

In order to build a regression model, a fuzzy inference system can be defined as a mapping between a vector of crisp inputs and a crisp output. Let us denote as scalar values $x_1, x_2, \ldots, x_M$ the inputs for observations $(x_k; y_k)$ of a certain regression problem. Assuming that all the inputs (M) are used, the fuzzy regression model can be expressed as a set of N fuzzy rules of the following form:

$$R_i\colon IF x_1 \; is \; L_1^i \wedge x_2 \; is \; L_2^i \wedge \ldots \wedge x_M \; is \; L_M^i \; THEN \; \mu_{R_i} \rightarrow \hat{y} \quad (2)$$

where $i=1,\ldots,N$, and the fuzzy sets $L_j^i \in \{L_{j,k}\}\, k = 1,\ldots,n_j, j = 1,\ldots,M$ where n_j is the number of linguistic labels defined for the j-th input variable. The L_j^i are the fuzzy sets representing the linguistic terms used for the j-th input in the i-th rule of the fuzzy model. The μ_{R_i} are the consequents of the rules and can take different forms. For example, in a system with two inputs, if L_1^i is renamed LOW_1 and L_2^i is renamed $HIGH_2$, the i-th rule R_i will have the following form:

$$R_i: IF\ x_1\ is\ LOW_1 \wedge x_2\ is\ HIGH_2\ THEN \mu_{R_i} \rightarrow \hat{y} \quad (3)$$

Depending on the fuzzy operators, inference model and type of membership functions (MFs) employed, the mapping between inputs and outputs can have different formulations. In principle, the methods proposed in this paper can be applied for any combination of types of MFs, operators and inference model, but the selection can have a significant impact on practical results.

As a concrete implementation for this paper, we use the minimum as T-norm for conjunction operations, Gaussian MFs for inputs, singleton outputs, and product inference of rules. Defuzzification is performed using the fuzzy mean method, i.e., zero-order Takagi-Sugeno systems (Nguyen & Prasad, editors., 1999) are defined. Thus, the result of the inference process is a weighted average of the singleton consequents. This inference scheme was chosen in order to keep systems as simple and interpretable as possible. In particular, the use of singleton outputs simplifies both the interpretation of rules and the local optimization process.

Therefore, in this particular case a fuzzy regressor can be formulated as follows:

$$F(\vec{x}) = \frac{\sum_{i=1}^{N} (\mu_{R_i} \cdot min_{1\le j\le M} \mu_{L_j^i}(x_j))}{\sum_{i=1}^{N} min_{i\le j\le M} \mu_{L_j^i}(x_j)} \quad (4)$$

where N is the number of rules in the rule base, the μ_{R_i} are singleton output values, and the $\mu_{L_j^i}$ are Gaussian MFs for the inputs. Thus, each fuzzy set $\mu_{L_j^i}$ (for the ith linguistic term defined for the jth input), is characterized by an MF having the following form:

$$\mu_{L_{j,k}} = \exp[-(y_j - c_{k,j})^2 / 2\sigma_{k,j}^2], k = 1, \dots, n_j, j = 1, \dots, M \quad (5)$$

where $c_{k,j}$ and $\sigma_{k,j}$ are scalar values and represent the centers and widths of the inputs MFs, respectively.

Fuzzy inference systems of the class being designed here are universal approximators (Sandín & Chorot, 1996; Jang et. al., 1997). Thus, for a sufficiently large number of rules and MFs, any input-output mapping can be approximated with arbitrary accuracy.

2.2.1.Clustering-Based Identification of Fuzzy Inference Systems

Different approaches to the identification of fuzzy inference systems from numeric data have been proposed in the literature (Mitra & Hayashi, 2000; Rutkowski, 2004). Roughly, two classes of methods can be distinguished: structure-oriented and clustering-based.

In this paper we focus on the clustering-based class of methods and especially on those methods that follow an offline approach. The following clustering algorithms are compared for the purposes of identifying fuzzy inference systems: The Hard and Fuzzy C-means (HCM and FCM, respectively) (Oliveira & Pedrycz, editors, 2007) clustering algorithms, the Improved Clustering for Function Approximation (ICFA) algorithm (Guillén et. al., 2007[1]), and the hybrid Fuzzy-Possibilistic Clustering for Function Approximation (Guillén, et. al., 2007[2]). The latter two algorithms were originally proposed for initializing Radial Basis Function Neural Networks (RBFNNs) for regression problems. In this paper, the $ICFA_f$ variant, tailored for fuzzy inference systems identification, is used (Pouzols & Barros, to appear).

The first step for clustering-based identification of fuzzy inference systems is to apply a clustering algorithm on the input-output patterns. Once this process finishes, Q clusters have been identified. The structure of the corresponding fuzzy inference systems then has to be defined. In general, fuzzy rules can be interpreted as joint constraints (Rutkowski, 2004) rather than implication rules. Thus, it is sensible to define a fuzzy rule from each cluster identified. This is the most frequent approach in the literature. This way, the clusters and their corresponding rules are considered as prototypes or models of the whole input pattern sequence.

Let us consider as above the case of a multiple scalar input and of a single scalar output where the input patterns entered into the clustering algorithm consist of M inputs and one output. Let us denote the clusters identified by $\vec{c}_k\ k = 1, \dots, Q$. Let every cluster have the following general form:

$$\vec{c}_k : (c_{k,1}, \dots, c_{k,M+1})\ with\ k = 1, \dots, Q \quad (6)$$

where the $c_{k,M+1}$ correspond to the outputs of the fuzzy inference model, whereas the $c_{k,1}, \dots, c_{k,M}$ correspond to the inputs $(x_1, \dots, x_M)$ of the fuzzy model. For each cluster, a matching rule is generated with the following form:

$$R_k: IF x_1\ is\ L_{1,k} \wedge x_2\ is\ L_{2,k} \wedge \dots \wedge x_M\ is\ L_{M,k}\ THEN\ c_{k,M+1} \rightarrow \hat{y}\quad k = 1, \dots, Q, Q = N = n_j \quad (7)$$

where a set of input linguistic terms is created $\{L_{j,k}\}, k =$

$1,\dots,n_j, j=1,\dots,M$. These linguistic terms are defined by Gaussian MFs $\mu_{L_{j,k}}$, as in equation (5). The output membership functions are defined as singleton functions centered at the corresponding element of the cluster centers $c_{k,M+1}$. The centers of the input Gaussian MFs for the jth input and kth rule ($c_{k,j}$ in equation (5)) are set to the jth elements of the corresponding clusters $\vec{c}_k$.

When inference systems are identified with clustering methods following this approach, the number n_j of linguistic terms defined for every input variable, $j = 1,\dots,M$, is equal to the number Q of clusters identified which in turn is equal to the number N of rules identified. Hence Q different membership functions are generated for each input and output variable, and Q rules are generated for a horizon h.

The way the widths of the input Gaussian MFs ($\sigma_{k,j}$ in equation (5)) are set depends on the clustering algorithm used. For the Hard C-means and Fuzzy C-means algorithms the widths are set as a function of the membership degrees of the input patterns to the clusters.
Recently, an adaptation of the ICFA (Guillén et. al., 2007[1]) algorithm for the identification of FIS was proposed (Pouzols & Barros, to appear). This adaptation, ICFA, is a simple generalization of the original ICFA proposal where all the widths for a certain rule are set to a value inversely proportional to the average weighting parameter w.

The ICFA algorithm performs an initialization of the centers of the clusters, taking into account the output of the function to be approximated. The output is considered by defining a value for each center in the output space. This value is named expected output (o_i) of a center i and allows the algorithm to weigh the distance computed between the input vectors and each center.

2.2.2. Fuzzy-Possibilistic approach

As was shown by Guillén et. al. (2007[1]), the combination of possibilistic and fuzzy membership functions could lead to a better center initialization for RBFNNs.

The development of the FPCFA algorithm relies on the approach presented by Pal et. al., (1997) where a combination of a fuzzy partition and a possibilistic partition is used. The authors assert that the membership value of the fuzzy partition is important in order to be able to assign a hard label to classify an input vector, but that at the same time, it is very useful to use the typicality (possibility) value to move the centers properly in the presence of outliers. Let $U^p = [u^p_{i,k}]$ be the matrix containing all the possibilistic memberships, $U^f = [u^f_{i,k}]$ the matrix containing the fuzzy memberships, and $C = [c_{i,k}]$ the matrix containing the center positions for $i=1\dots m$ and $k=1\dots n$. The distortion function to be minimized is:

$$J_{h_f,h_p}(U^f, C, U^p; X) = \sum_{k=1}^{n}\sum_{i=1}^{m}((u^f_{ik})^{h_f} + (u^p_{ik})^{h_p})D^2_{ik} \quad (8)$$

with the following constraints:

$$\begin{aligned} &\sum_{i=1}^{m} u^f_{ik} = 1 \forall k = 1 \dots \mathrm{n} \\ &\sum_{k=1}^{n} u^p_{ik} = 1 \forall i = 1 \dots \mathrm{m} \end{aligned} \quad (9)$$

The constraint shown above requires each row of U^p to sum up to 1 but its columns are free up to the requirement that each column contains at least one non-zero entry. Therefore, there is a possibility of input vectors not belonging to any cluster. The design of the FPCFA algorithm weighs the similarity criteria used in the computation of the distances and defines an expected output for each center, so the distortion function to be optimized remains:

$$J_{h_f,h_p}(U^f, C, U^p; X) = \sum_{k=1}^{n}\sum_{i=1}^{m}((u^f_{ik})^{h_f} + (u^p_{ik})^{h_p})D^2_{ikW} \quad (10)$$

restricted to the same constraints as for the FPCM algorithm.

The iteration method used for minimization considered the following equations to compute the membership and expected output:

$$\begin{aligned} u^f_{ik} &= \sum_{j=1}^{m}\left(\frac{D_{ikW}}{D_{jkW}}\right)^{\frac{2}{h_f-1}^{-1}} \\ u^p_{ik} &= \sum_{j=1}^{n}\left(\frac{D_{ikW}}{D_{ijW}}\right)^{\frac{2}{h_p-1}^{-1}} \\ \vec{c}_k &= \frac{\sum_{k=1}^{n}((u^f_{ik})^{h_f}+(u^p_{ik})^{h_p})\vec{x}_k w^2_{ki}}{\sum_{k=1}^{n}((u^f_{ik})^{h_f}+(u^p_{ik})^{h_p})w^2_{ki}} \\ o_i &= \frac{\sum_{k=1}^{n}((u^f_{ik})^{h_f}+(u^p_{ik})^{h_p})y_k d^2_{ki}}{\sum_{k=1}^{n}((u^f_{ik})^{h_f}+(u^p_{ik})^{h_p})d^2_{ki}} \end{aligned} \quad (11)$$

The algorithm iterates until the centers have not moved significantly.

3. Experiments

This section will demonstrate the results obtained after applying the methods described in the previous section. First, variable selection will be performed in order to reduce the dimensionality and obtain a simpler design, and to identify key risk factors. Afterwards, RBFNN will be implemented to classify the data set.

The data set consists of 5452 cases of 22 variables; 1788 cases were randomly selected as the test sample and 3664 as the training sample. The output variable is the number of weeks required to process the dissolution. These records were taken from the first data base available for all verdicts on marital dissolutions in Spain during the year 2007. The database was provided by the Spanish National Institute of Statistics (INE, Instituto Nacional de Estadística) and has been used for research in the Department of Preventive Medicine of the University of Granada.

3.1.Dimensionality Reduction

As described in the previous section, the Mutual Information (MI) will be used in order to rank the variables. Since it is not possible to compute the MI for all the possible combinations of input variables and the output, a compromise approach was adopted which consisted in computing the MI for each variable and the output. The results provided are shown in Table 1 and graphically depicted in Figure 1.

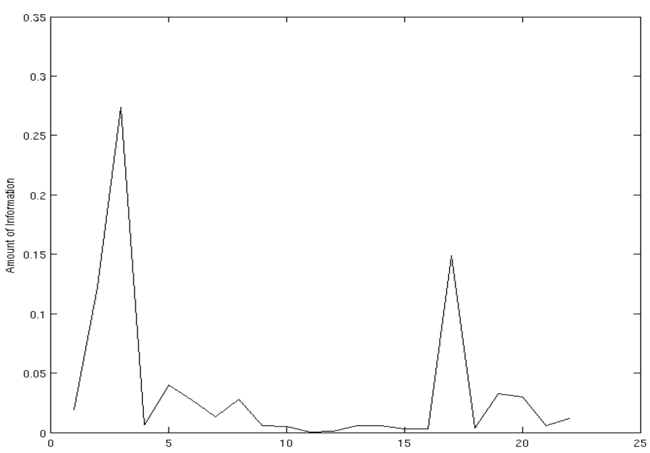

Figure 1. Mutual Information of predictors for the training data set.

Parsimonious models can now be established by combining classical assumptions within each mixture on both mixing proportions and *t*-parameters (intrapopulation models), with meaningful constraints on the parametric link (3) between the conditional populations (interpopulation models).

Table 1. Mutual information (MI) for each variable of the training data set.

MI value	Var number and description
0.001	11 – Husband's previous civil state
0.001	12 – Wife's Birth Month
0.003	16 – Wife's previous civil state
0.003	15 – Wife's Nationality
0.004	18 – Monthly pay
0.005	10 – Husband's Nationality
0.006	21 – Previous Separation
0.006	9 – Sex of the first couple (husband)
0.006	14 – Sex of the second couple (wife)
0.006	13 – Wife's Birth Year
0.006	4 – Month of Marriage
0.012	22 – Region
0.013	7 – Husband's Birth Month
0.019	1 – Province
0.027	6 – Number of underage children
0.028	8 – Husband's birth year
0.03	20 – Custody
0.033	19 – Alimony
0.04	5 – Year of Marriage
0.123	2 – Month of Lawsuit
0.149	17 – Claimant
0.27	3 – Year of the Lawsuit

3.2.Dissolution Duration Prediction

After ranking the variables, it is possible to start designing models to predict the duration of the process in weeks.

Table 2. Test results for predictor 3 (Year of lawsuit). Average and standard deviation of the test RMSE (normalized)

# clusters	HCM	FCM	ICFA	FPCFA
2	.701 ± .490	0.701 ±0.490	0.701 ±0.490	1.000 ±0.758
3	.701 ± .489	0.857 ±0.647	0.857 ±0.647	0.700 ±0.489
4	.700 ± .489	0.700 ±0.489	0.700 ±0.489	0.700 ±0.489
5	.700 ± .489	0.700 ±0.489	0.701 ±0.490	0.700 ±0.489
6	.700 ± .489	0.700 ±0.489	0.701 ±0.489	0.700 ±0.489
7	.700 ± .489	0.700 ±0.489	0.701 ±0.489	0.700 ±0.489
8	.700 ± .489	0.702 ±0.490	0.700 ±0.489	0.700 ±0.489
9	.700 ± .481	0.700 ±0.489	0.701 ±0.489	0.700 ±0.489
10	700 ± .481	0.702 ±0.490	0.701 ±0.490	0.700 ±0.489
15	.700 ± .489	0.700 ±0.489	0.700 ±0.489	0.700 ±0.489
20	.700 ± .489	0.700 ±0.489	0.700 ±0.489	0.700 ±0.489

Table 3. Test results for predictors 3 and 17 (claimant). Average and standard deviation of the test RMSE (normalized)

# clusters	HCM	FCM	ICFA	FPCFA
2	0.676 ±0.481	0.675 ±0.480	0.673 ±0.479	0.672 ±0.479
3	0.650 ±0.464	0.662 ±0.476	0.655 ±0.471	0.649 ±0.464
4	0.659 ±0.474	0.787 ±0.604	0.681 ±0.484	0.650 ±0.464
5	0.716 ±0.532	0.659 ±0.474	0.647 ±0.462	0.650 ±0.464
6	0.659 ±0.474	0.648 ±0.463	0.648 ±0.462	0.648 ±0.462
7	0.654 ±0.470	0.649 ±0.464	0.651 ±0.465	0.648 ±0.462
8	0.655 ±0.469	0.659 ±0.473	0.649 ±0.463	0.648 ±0.463
9	0.649 ±0.464	0.650 ±0.464	0.647 ±0.462	0.648 ±0.463
10	0.654 ±0.470	0.657 ±0.472	0.649 ±0.464	0.668 ±0.489
15	0.649 ±0.464	0.650 ±0.465	0.648 ±0.462	0.649 ±0.463
20	0.649 ±0.464	0.649 ±0.464	0.649 ±0.463	0.648 ±0.462

Table 4. Test results for predictors 3, 17 and 2 (month of lawsuit). Average and standard deviation of the test RMSE (normalized)

# clusters	HCM	FCM	ICFA	FPCFA
2	0.587 ±0.399	0.566 ±0.399	0.619 ±0.455	0.568 ±0.398
3	0.618 ±0.448	0.551 ±0.392	0.537 ±0.380	0.546 ±0.390
4	0.727 ±0.574	0.569 ±0.409	0.521 ±0.366	0.513 ±0.360
5	0.668 ±0.520	0.815 ±0.645	0.550 ±0.395	0.511 ±0.362
6	0.542 ±0.392	0.726 ±0.565	0.541 ±0.388	0.511 ±0.365
7	0.517 ±0.364	0.723 ±0.563	0.507 ±0.360	0.510 ±0.364
8	0.508 ±0.356	0.670 ±0.482	0.511 ±0.363	0.512 ±0.366
9	0.508 ±0.367	0.523 ±0.373	0.501 ±0.355	0.506 ±0.361
10	0.506 ±0.360	0.504 ±0.360	0.511 ±0.365	0.610 ±0.466
15	0.506 ±0.361	0.506 ±0.359	0.505 ±0.360	0.512 ±0.370
20	0.503 ±0.361	0.514 ±0.373	0.503 ±0.359	0.505 ±0.361

Table 5. Test results for predictors 3, 17, 2 and 5 (year of marriage). Average and standard deviation of the test RMSE (normalized)

# clusters	HCM	FCM	ICFA	FPCFA
2	0.566 ±0.401	0.565 ±0.399	0.623 ±0.453	0.558 ±0.390
3	0.822 ±0.646	0.804 ±0.635	0.598 ±0.438	0.542 ±0.382
4	0.534 ±0.385	0.571 ±0.422	0.526 ±0.375	0.589 ±0.431
5	0.707 ±0.546	0.544 ±0.395	0.511 ±0.364	0.513 ±0.357
6	0.578 ±0.419	0.532 ±0.380	0.545 ±0.392	0.534 ±0.374
7	0.566 ±0.420	0.714 ±0.559	0.536 ±0.387	0.505 ±0.358
8	0.534 ±0.388	0.915 ±0.727	0.569 ±0.409	0.512 ±0.361
9	0.526 ±0.378	0.633 ±0.470	0.509 ±0.364	0.514 ±0.367
10	0.512 ±0.364	0.539 ±0.388	0.509 ±0.359	0.512 ±0.367
15	0.525 ±0.380	0.517 ±0.371	0.508 ±0.363	0.511 ±0.368
20	0.510 ±0.362	0.522 ±0.378	0.503 ±0.359	0.520 ±0.375

The procedure followed was incremental: the variable with the highest MI value was chosen, then, the first and the second, and so on. Tables 2-5 present the results obtained after the execution.

3.3. Results Discussion

The most relevant variable was the year of lawsuit, which is directly influenced by the increase in the computerization of the bureaucracy involved, which accelerated the process. It follows that recent divorces take less time. Unfortunately, the variable on extent of computerization was not been measured by the data providers.

Regarding temporal variables, the month in which the lawsuit starts is very relevant. The cause might be the coincidence with holiday periods (summer and Christmas), where the staff might work with less productivity, resulting in longer divorces.

The fourth variable, year of marriage, also adds interesting information. As the data show, old marriages and young marriages take less time to divorce. A deeper analysis of the reasons for this fact might fall out of the scope of the paper.

Looking at the results from the machine learning perspective, it can be concluded that specific clustering techniques such as ICFA and FPCFA outperform the classical approaches HCM and FCM.

Figure 2 displays the minimum RMS test error obtained for each algorithm using the three most relevant variables. The approximations made by FPCFA and $ICFA_f$ are better for a small number of rules although all the algorithms have a very similar performance as the number of rules increases. However, it is important to obtain accurate results with a small number of rules in order to maintain the interpretability of the system.

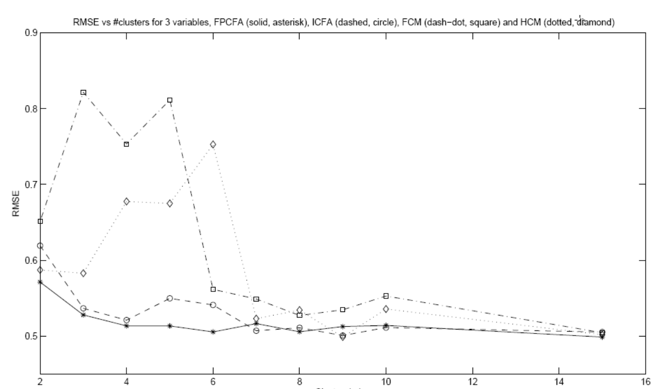

Figure 2. Comparison of RMS test errors for several numbers of rules

The experiments also show that the accuracy does not improve significantly as more variables in the current set of predictors are considered, indicating that for further studies, other variables should be sampled. Figure 3 depicts the minimum error obtained by the FPCFA in terms of the number of variables used. As the dimensionality increases over 3 variables, the accuracy starts to decrease.

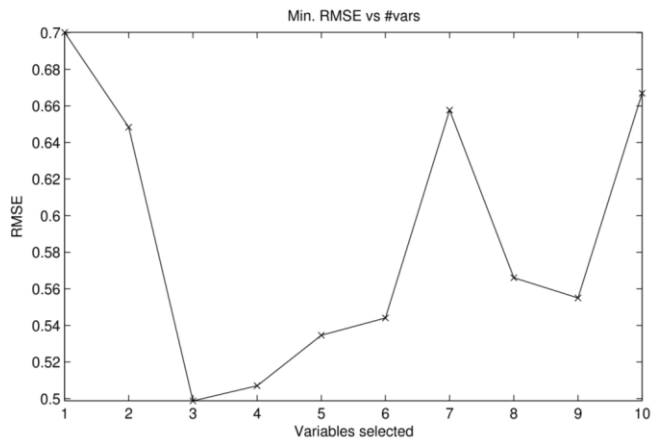

Figure 3. Minimum RMS test error obtained using several number of variables

4. Conclusions

The possibility of having models that are able to approximate data is quite useful in the field of preventive medicine. This paper has presented an example of such a model: fuzzy inference systems were used to predict the length of a marital dissolution process. The prediction could be used to prevent mental distress in the people involved in the dissolution (wife, husband and children) as well as for the assignation of human resource in administration. The methodology proposed here performed a previous step of dimensionality reduction using the mutual information concept which allows the models to provide better accuracy in the predictions.

References

European Commission. 2009. The state of mental health in the european union. In Directorate General for Health and Consumer Protection.

Fariña, F. Arce, R. Seijo, D. & Novo, M. 2003. Comparación de los niveles de adaptación entre menores pertenecientes a familias intactas y familias con padres separados. In IV Congreso Virtual de Psiquiatría Interpsiquis.

Guillén, A. González, J. Rojas, I. Pomares, H. Herrera, L. J. Valenzuela, O. & Prieto, A. 2007[1]. Using fuzzy logic to improve a clustering technique for function approximation. *Neurocomputing,* 70(16-18):2853 – 2860.

Guillén, A. Pomares, H. Rojas, I. González, J. Herrera, L.J. Rojas, F. & Valenzuela, O. 2007[2]. Studying possibility in a clustering algorithm for RBFNN design for function approximation. *Neural Computing & Applications,* 17(1):75 – 89.

Harald, S. Alexander, K. Sergey, A.A. Peter, G. 2004. Least dependent component analysis based on mutual information, *Phys. Rev. E,* 70.

Haro, J.M. Palacín, C. Vilagut, G. & Martínez, M. et. al. 2006. Prevalencia de los trastornos mentales y factores asociados: resultados del estudio esemed-españa. *Medicina Clínica,* 126:445 – 451.

Holmes, T.S. & Rahe, R.H. 1967. The social readjustment rating scale. *Journal Psychosomatic Research,* 11:213 – 218.

Hughes, M.E. & Waite, L.J. 2009. Marital biography and health at mid-life. *Journal of Health and Social Behavior,* 50:344 – 358.

Jang, J.S. Sun, C.T. & Mizutani, E. 1997. Neuro-Fuzzy and Soft Computing A Computational Approach to Learning and Machine Intelligence. Prentice Hall.

Kraskov, A. Stögbauer, H. & Grassberger, P. 2004. Estimating mutual information. *Phys. Rev. E*, 69(6).

Lansford, J.E. 2009. Parental divorce and children's adjustment. *Perspectives on Psychological Science,* 4(2):140 – 52.

Marí-Klose, P. Marí-Klose, M. Granados, F.J. Gómez-Granell, C. & Martínez. A. 2009. Informe de la inclusión social en españ,a. Fundación Caixa Cataluña.

Mitra, S. & Hayashi, Y. 2000. Neuro-fuzzy rule generation: survey in soft computing framework. *IEEE Transactions on Neural Networks,* 11(3):748 – 768.

Nguyen, H.T. & Prasad, N.R. editors. 1999. Fuzzy Modeling and Control: Selected Works of M. Sugeno. CRC Press, Inc.

Oliveira, J. V. & Pedrycz, W. editors. 2007. Advances in Fuzzy Clustering and its Applications. John Wiley & Sons, Ltd.

Organización Panamericana de la Salud. 2003. Familia y salud consejo directivo. 44.

Orgilés, M. Espada, J.P. & Méndez, X. 2008. Trastorno de ansiedad por separación en hijos de padres divorciados. *Psicothema,* 20:383 – 388.

Pal, N.R. Pal, K. & Bezdek, J.C. 1997. A mixed c-means clustering model. *In Proceedings of the 6th IEEE International Conference on Fuzzy Systems (FUZZ-IEEE'97),* 1:11 – 21.

Pouzols, F.M. & Barros, A.B. Automatic clustering-based identification of autoregressive fuzzy inference models for time series. *Neurocomputing,* to appear.

Rutkowski, L. 2004. Flexible Neuro-Fuzzy Systems. Structures, Learning and Performance Evaluation. Kluwer Academic Publishers.

Sandín, B. & Chorot, P. 1996. Evaluación del estrés psicosocial. *Manual de evaluación en psicología clínica y de la salud,* 619 – 647.

World Health Organization. 2005. Public health action for healthier children and populations. Regional Office for Europe The European health report 2005.

Ying, H. 1998. Sufficient Conditions on Uniform Approximation of Multivariate Functions by General Takagi-Sugeno Fuzzy Systems with Linear Rule Consequent. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans,* 28(4):515 – 520.

Acknowledgements F. Montensino is supported by a Marie Curie Intra-European Fellowship for Career Development (grant agreement PIEF-GA-2009-237450) within the European Community's Seventh Framework Programme (FP7/2007--2013). This work was partially supported by the Spanish CICYT Projects TIN2007-60587 and TEC2008-04920 and Junta Andalucia Projects P07-TIC-02768, P08-TIC-03674 and P08-TIC03928, and PYR-2010-17 of CEI Bio GENIL (CEB09-0010) of MICINN. The authors also wish to thank the Spanish Institute of Statistics for providing the data.

Correspondence: mailto:aguillen@atc.ugr.es