Revista Internacional de
Sostenibilidad, Tecnología y Humanismo
año 2006 nº 1

Connectionist techniques to approach sustainability modelling

Marta Domingo*, Núria Agell** and Xavier Parra***

*GREC Group and EPSEVG-UPC. Av. Víctor Balaguer s/n. 08800 Vilanova i la Geltrú, Spain. marta.domingo-gou@upc.edu
**GREC Group and ESADE-URL. Av. Pedralbes, 60. 08034 Barcelona, Spain. nuria.agell@esade.edu
***GREC Group and ESAII-UPC. Av. Víctor Balaguer s/n. 08800 Vilanova i la Geltrú, Spain. xavier.parra@upc.edu
Web page: http://www.upc.edu/web/grec

Summary

When defining a context of sustainability, capturing the complexity of data and extracting as much information as possible are fundamental challenges. Normally, quantitative and qualitative indicators are defined. While the definition and calculation of the former is direct, the latter are difficult to manage. This document provides tools based on connectionist techniques for managing complex information combining the use of imprecise and qualitative variables. The application of these tools to evaluate non-numerical sustainability indicators is presented. The results obtained in some first approaches are briefly presented to illustrate the connectionist paradigm.

Keywords: sustainability, sustainability modelling, sustainable development valuation, learning, connectionist techniques, citizen's satisfaction.

1 Introduction

This paper presents work in the field of Artificial Intelligence applied to sustainability modelling. When defining a context of sustainability, capturing the complexity of data and extracting as much information as possible are fundamental challenges. Normally, quantitative and qualitative indicators are defined. While the definition and calculation of the former is direct, the latter are difficult to manage. Our principal concern is to analyze the interest of applying AI to the sustainability measuring topic. These are adaptations of connectionist techniques which are capable of managing qualitative concepts as well as quantitative data to make progress in modelling complex systems. Specifically, we are exploring forms of Artificial Neural Networks, Fuzzy Systems & Hybrid Connectives, Evolutionary Computation and Kernel Methods.

The definition of indicators is a common approach to monitoring sustainability. This is the case of the European Common Indicators initiative, which is focused on monitoring sustainability at the local level. The European Common Indicators (2006) are a ready-to-use, self-contained set of 10 indicators that help a town or city, interested in the quality of its urban environment, to monitor progress. Our recent work concerns the qualitative indicator 'Citizen Satisfaction with the Local Community', which was the first of the set.

Our research group (Ferrer, Català and Angulo, 2002) was commissioned with a descriptive study of this indicator by the town council of Vilanova i la Geltrú (Catalonia, Spain), in which data were collected through a questionnaire focused on citizens. Using a system capable of learning from qualitative data we have explored the possibility of modelling citizens' satisfaction including some tacit knowledge revealed in the citizens' answers (Domingo et al., 2004), (Domingo et al., 2005). Our future research aims to improve this kind of model to predict the evolution of citizens' satisfaction under certain input changes. This would provide local councils with a tool to support decision making.

2 Why a connectionist approach?

The design of algorithms able to automatically gather the relevant information from a set of patterns is one of the primary aims of Artificial Intelligence. In recent years some artificial intelligence techniques have shown to be useful to modelize complex problems. On the one hand, problems in which it is necessary to deal with ambiguous or uncertain information have a proper representation when using connectionist techniques. On the other hand, those that are related to common sense reasoning have a proper representation when using qualitative variables. The main goal is to obtain models that are adaptable to a given application. In the application we are considering both conditions are fulfilled, there is a high implication of ambiguous variables and the use of common sense reasoning is directly involved in the questionnaire variables.

In general, the use of learning and connectionist algorithms permit the adaptation to situations in which statistical and data mining tools are too rigid.

3 Computational intelligence Techniques

There is a wide variety of methods available for the classification of data, in general, and for analysis in particular. These methods are derived from techniques used in different domains like statistics, clustering, and others. A group of alternative approaches is based on computational intelligence techniques and, among others, include artificial neural networks, fuzzy systems, evolutionary computation and kernel methods.

3.1 Artificial Neural Networks

An artificial neural network (ANN) can be characterized as a computational model with particular properties, such as the ability to learn, to adapt and to generalize. The emergence of the early architectures for neural networks is biologically inspired, and the basic premise of this alternative approach is that such systems perform extraordinarily complex computations in the real world without recourse to explicit quantitative operations and can process complex data in a flexible manner, that is relatively independent of the task defined.

An ANN is an interconnected group of artificial neurons that uses a mathematical or computational model for information processing based on a connectionist approach to computation. In more practical terms neural networks are non-linear statistical data modelling tools that can be used to model complex relationships between inputs and outputs or to find patterns in data.

Typically, a neural network consists of many computational elements or nodes arranged in layers. They are specified by the network architecture, node characteristics and learning algorithms. There are many different types of neural networks, but probably the most well known are the feedforward neural networks. A feedforward neural network consists of a set of nonlinear neurons connected together, in which the information flows only in the forward direction, from inputs through the hidden nodes (if any) and to the outputs (i.e. there are no cycles or loops in the network). Probably, the most widely used feedforward networks are: Multilayer Perceptron (MLP) and Radial Basis Function Networks (RBF).

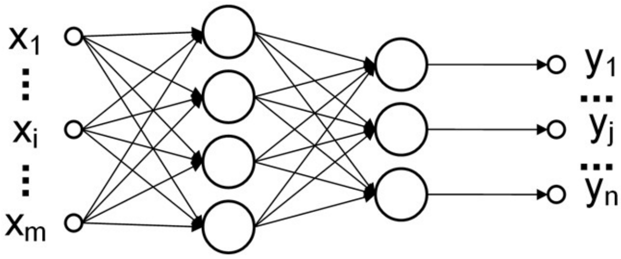

Figure 1: A Multilayer Perceptron with m inputs, a hidden layer and n outputs

A Multilayer Perceptron consists of a set of input nodes, an output neural layer, and a number of layers of hidden neural nodes, an example of which can be shown in Figure 1. If the input variables of an MLP take continuous values in [0,1], one layer of artificial neurons builds a discriminative hyperplane in the input space; the second layer organizes these discrimination hyperplanes, so that arbitrarily complex convex separation surfaces can be approximated. A third neural layer would allow the definition of also non convex and disjoint separation surfaces in the input space.

MLPs are universal approximators (Hornik, Stinchcombe and White, 1989) and therefore universal classifiers. This means that if the activation function is the logistic sigmoid, it can be proven that any continuous decision boundary – and therefore any generally smooth mapping – can be approximated arbitrarily close by a two-layer perceptron with a sufficient number of hidden neurons. Thus MLPs provide universal nonlinear discriminant functions.

RBF networks are especially interesting for classification problems since they are universal classifiers (Poggio and Girosi, 1990). RBFs make use of local receptive fields to associate and to establish correspondences between inputs and outputs and have been traditionally associated with a simple architecture of three layers (Broomhead and Lowe, 1988) (see Figure 2). Each layer is fully connected to the following one. The hidden layer is composed of a number of nodes with radial activation functions called radial basis functions. Each of the components of the input vector feeds forward to the basis functions whose outputs are linearly combined with weights into the network output. Each radial function has a local response (opposite to the global response of sigmoid function) since their output only depends on the distance of the input from a centre point.

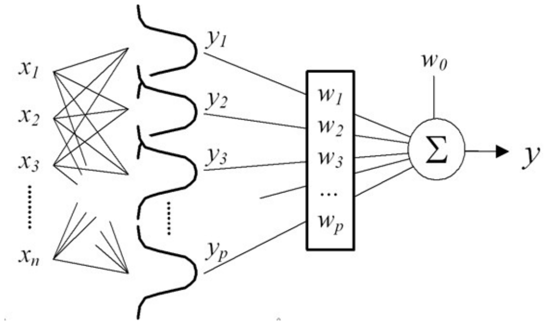

Figure 2: A Radial basis function network architecture.

Radial functions in the first layer have a special structure that can be represented as follows:

$$y = f\left(\left\| x - c_i \right\| \right) \tag{1}$$

where C_i is the radial function centre and || · || typically denotes the Euclidean norm. The traditional radial function used with RBF networks is the Gaussian function, which has an additional parameter such as its width (σ). The function associated to a Gaussian radial basis function network is defined as:

$$f(x) = w_0 + \sum_{i=1}^{p} w_i e^{\left(-\left\|x - c_i\right\|^2 / \sigma_i^2\right)} \tag{2}$$

where p is the number of basis functions and w_i are the synaptic weights of the output layer.

3.2 Fuzzy Systems and Hybrid Connectives

In most real-world scenarios there is always a certain degree of uncertainty in the data, i.e. data will not be completely precise. This uncertainty can be due to the measurement process (e.g. control application) or to the nature of the problem (e.g. questionnaire with closed options for answers). Several approaches to handle information about uncertainty have already been proposed, for example interval arithmetic allows us to deal and compute with intervals rather than crisp numbers, and also numerical analysis offers ways to propagate errors along with the normal computation. However, one of the techniques most used with imprecise concepts are those based on *fuzzy logic* (Zadeh, 1965). This type of logic enables us to handle uncertainty in a very intuitive and natural manner. In addition to making it possible to formalize imprecise numbers, it also enables us to do arithmetic using such *fuzzy numbers*. Classical set theory can be extended to handle partial memberships, thus making it possible to express vague human concepts using *fuzzy sets* and also describe the corresponding inference systems based on *fuzzy rules*.

The general scheme of a fuzzy system can be seen in Figure 3. Basically, the system is structured in three parts: fuzzification, inference and defuzzification.

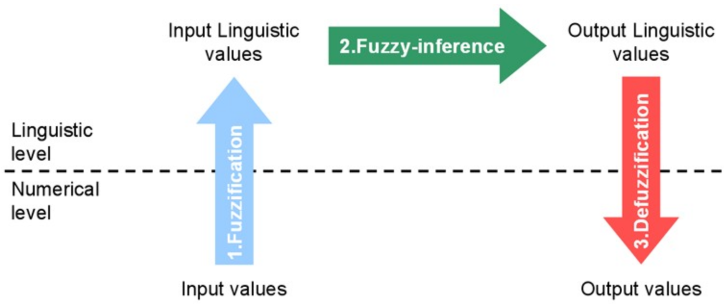

Figure 3: Fuzzy system scheme

A classifying technique called LAMDA (Learning Algorithm for Multivariate Data Analysis) based on hybrid connectives can be used. The LAMDA classifier relies on the generalizing power of Fuzzy Logic and the interpolation capability of logical hybrid connectives. One of its main advantages is the capability to cope simultaneously with numerical and qualitative information. The implementation of these possibilities takes the form of an algorithm developed by Josep Aguilar in collaboration with other authors (Aguado et al., 1999), who have been enhancing this original self-learning classifying technique since the eighties. LAMDA can perform either supervised or unsupervised learning, i.e. it can either mimic the classification criteria of a human expert or it can automatically search for logical data segmentations.

The adequacy of an individual to a class, such as a group of citizens who show a similar behaviour is defined as the combination of the attraction of each of the partial descriptors Xi that define the profile X with respect to the corresponding component of xi of the individual x. In LAMDA algorithm it is assumed that a Marginal Adequacy Degree, MAD(xi/Xi) can be evaluated for each descriptor, and they are synthesized by using a hybrid connective. This connective L is defined as a linearly compensated combination of two fuzzy operators (Klir and Yuan, 1995), a t-norm and its dual t-conorm and permits to obtain the Global Adequacy Degree of an individual x to a class or profile X.

$$GAD(x/X)=L(MAD(x_1/X_1),\ MAD(x_2/X_2),\ldots,\ MAD(x_n/X_n)) \quad (3)$$

In the above the L is defined by means of a parameter λ that will determine the exigency level of the classification: L= λ·T+(1-λ)·T*, where T is a t-norm (for example Min) and T* is its dual t-conorm (for example Max).

3.3 Evolutionary Computation

Evolution in nature provides fascinating mechanisms for adapting living species to their environments. Trying to integrate evolution into algorithmic schemes, propagating the field known as *Evolutionary Computation*, has been an emerging area of development in artificial intelligence. Probably one of the most well known methodologies in this field is *Genetic Algorithms* (GA).

Genetic Algorithms have been popularized as universal optimization algorithms based on evolutionary methods (Holland, 1975). GAs belongs to the group of stochastic optimization methodologies. These stochastic methods turn the search of optimal solution towards a zone where one hopes to find solution and allow that the randomness helps to find good parameters. Normally, GAs use a binary parameter encoding resembling the discrete nucleotide coding scheme on cellular chromosomes. Therefore, GA genotypes can be defined as bit-vectors on which point mutations are defined by switching bits with a certain probability. A recombinational operator of crossover models the breaking of two chromosomes during reproduction. Each individual is assigned a fitness value depending on the optimization task to be solved. Selection within the population is performed in a fitness-proportionate way. The more fit an individual, the more likely it is to be chosen for reproduction into the following generation.

There are four main elements in GA: population (group of individuals), genetic operators (source of variation), fitness (reproductive fitness) and selection (survival of the fittest). The evolutionary cycle consists on creating a population (diverse and different machines) and scoring the adjustment or adaptation (fitness) of each individual on a version of the task to develop (e.g.

success on a training set). Afterwards, the individuals are ordered (rating) based on its adaptation and the best adapted survive. Finally, stochastic alteration is applied to complete the following generation (i.e. recombination and mutation).

3.4 Kernel Methods

In the last years, a number of powerful kernel-based learning machines, e.g. Support Vector Machines (SVM), Kernel Fisher Discriminant (KFD) or Kernel Principal Component Analysis (KPCA), have been proposed and showed practical relevance for classification and regression problems.

SVMs are learning systems that use a hypothesis space of linear functions in a high dimensional feature space, trained with a learning algorithm from optimisation theory that implements a learning bias derived from statistical learning theory. This learning strategy introduced by Vapnik (Vapnik, 1995) is a principled and very powerful method that in the few years since its introduction has already outperformed most other systems in a wide variety of applications. SVMs use a technique known as the *kernel trick* to apply linear classification techniques to non-linear classification problems.

In machine learning, the *kernel trick* is a method for easily converting a linear classification learning algorithm into a non-linear one, by mapping the original observations into a higher-dimensional non-linear space so that linear classification in the new space is equivalent to non-linear classification in the original space. This is done using Mercer's condition, which states that any positive semi-definite kernel K(x, y) can be expressed as a dot product in a high-dimensional space. More specifically, if the arguments to the kernel are in a measurable space X, and if the kernel is positive semi-definite — i.e.

$$\sum_{i,j} K(x_i, x_j) c_i c_j \geq 0 \qquad (4)$$

for any finite subset $\{x_1, ..., x_n\}$ of X and subset $\{c_1, ..., c_n\}$ of real numbers — then there exists a function φ(x) whose range is in an inner product space of possibly high dimension, such that

$$K(x,y) = \varphi(x) \cdot \varphi(y) \qquad (5)$$

The kernel trick transforms any algorithm that solely depends on the dot product between two vectors. Wherever a dot product is used, it is replaced with the kernel function. Thus, a linear algorithm can easily be transformed into a non-linear algorithm. This non-linear algorithm is equivalent to the linear algorithm operating in the range space of φ. However, because kernels are used, the φ function is never explicitly computed. This is desirable, because the high-dimensional space may be infinite-dimensional (as is the case when the kernel is a Gaussian).

4 Problem presentation

In the European context, a definition of sustainability indicators was addressed at the meeting: 'Towards a Local Sustainability Profile - European Common Indicators' during the '3rd European Conference on Sustainable Cities and Towns' (9-12 February 2000, Hanover, Germany). At this meeting, local authorities of many European cities agreed to integrate the European Common Indicators into their existing municipal management systems. This initiative is a step towards a new generation of monitoring practices.

The monitoring initiative is designed to support local authorities in their work towards sustainability by measuring movement towards or away from sustainability, and focusing on the extent of change over time and the identification of trends and directions, more than on absolute measures.

A set of environmental sustainability indicators and methodologies for collecting the data for each indicator was developed in conjunction with stakeholders. A final list of 10 integrated indicators was obtained through a process of analysis and elimination of many existing indicators.

The resulting 10 European Common Indicators were:

– Citizen satisfaction with the local community

– Local contribution to global climatic change

– Local mobility and passenger transportation

– Availability of local public open areas and services

– Quality of local ambient air

– Journeys by children to and from school

– Sustainable management of the local authority and local business

– Noise pollution

– Sustainable land use

– Products promoting sustainability

The definition and calculation of some indicators is direct, because they are quantitatively measurable (in terms of CO_2 emissions, passenger transport (km/capita), percentage of population with access to green spaces, number of days with good air quality, etc.). Other indicators, such as the 'degree of satisfaction of citizens with their locality', involve a qualitative perception of reality and have to be managed differently.

In other words, the economic, social and environmental systems that make up the community are partial issues that may give an idea of the quality of life in that community, which is the central concern of sustainability. However, the satisfaction of citizens with their local community is about more than just numerical accordance of their locality with permitted ratings.

In many works, quantitative indicators are taken into account as a central concern of urban sustainability, while other subjective perceptions of people related to global welfare are not properly dealt with. Here, we concentrate on the management of this subjective information.

5 On going research

5.1 Vilanova i la Geltrú, a case study

Vilanova i la Geltrú (Catalonia, Spain) is a town with over 55.000 inhabitants. In the Hanover meeting (February, 2000), the local authorities of Vilanova, together with many other European cities, took the compromise of integrating the European Common Indicators into the existing municipal management system. The 'degree of satisfaction of citizens with their locality' was chosen as the first Common Indicator. A descriptive study on the global subjective indicator 'citizens' satisfaction' in Vilanova i la Geltrú was carried out from December 2001 to November 2002 (Ferrer, Català and Angulo, 2002).

5.2 Framework and involved variables

The methodology proposed to study the 'citizens' satisfaction' indicator was based on a standard questionnaire provided by the European Common Indicators initiative that was composed of 11 questions about several aspects of the urban environment. There were 10 variables involved with the basic urban services (natural resources protection, carefulness of public spaces, employment, cultural and leisure events, health services, education services, public transport, local governance (planning and decision making processes), public safety services, housing). An 11th variable was a more synthetic and subjective issue dealing with the convenience of the city for everyone as a good place to live and work in (see Questionnaire below).

Questionnaire

1. Are you satisfied about the protection of natural resources (urban environment, beach, mountains)?
2. Are you satisfied about the streets, public spaces, façades, pedestrian-only zones?
3. Are you satisfied about the employment opportunities in Vilanova i la Geltrú?
4. Are you satisfied about the level of cultural, sports, and leisure services in Vilanova i la Geltrú?
5. Are you satisfied about the level of public health and social services in Vilanova i la Geltrú?
6. Are you satisfied about the education services in Vilanova i la Geltrú?
7. Are you satisfied about the public transport services in Vilanova i la Geltrú?
8. Are you satisfied about the opportunities for participation in the decision-making processes

of the town (municipal elections, forums, attention to citizens' problems, etc.)

9. Are you satisfied about the level of security in Vilanova i la Geltrú?

10. Are you satisfied about accommodation opportunities in Vilanova i la Geltrú (easiness, quality, prices)?

11. Are you satisfied about Vilanova i la Geltrú as a globally good place to live and work in?

The questionnaire provided each question with five possible linguistic labels as answers, expressing a given degree of satisfaction: Don't know, Very dissatisfied, Rather dissatisfied, Quite satisfied, Very satisfied.

Only minor changes and additions to the questionnaire were allowed to let the council achieve some qualitative subjective information. A 12th question was added to the frame questionnaire, asking for the reasons for satisfaction and dissatisfaction. This was an open question, that is, a direct qualitative answer was allowed in order to collect more detailed answers concerning the best-considered advantages of the town. The questionnaire was administered to a sample of 1000 citizens from Vilanova i la Geltrú, and a classical statistical study was performed.

There were two reasons that motivated us to go further. First, the indicator 'degree of satisfaction of citizens with their locality' has a strong qualitative nature. Second, the fact that the percentage of citizens who were quite satisfied or very satisfied with the town as a good place to live and work in (variable 11) was perceptibly higher than the percentage of quite satisfied or very satisfied persons in the partial questions (variables 1 to 10). This would suggest that part of the subjective information was missed out within the classical approach.

5.3 Results

Our research on the adaptation and application of a connectionist approach to the analysis of non-numerical sustainability indicators is a work in progress. Here, we briefly describe some experiments on this.

An automatic classifier algorithm called LAMDA (Logical Association for Multivariate Data Analysis) (Aguado, 1998) was applied to the analysis of a database dealing with the first indicator 'degree of satisfaction of citizens with their locality' (Domingo et al., 2004). Through a supervised learning process our objective was to identify and distinguish citizen profiles according to the satisfaction with the city they were living in.

Thus, urban management policies could be directed to improve the aspects that are more important to satisfy a given profile of citizen.

Moreover, we considered that it was possible to improve the initial learning process, taking advantage of the whole amount of pure qualitative information about the reasons of citizens' satisfaction and dissatisfaction collected in the open answer of the 12th question. Nevertheless, the collection and treatment of this kind of answers require considerable effort.

In a fist step, we defined different citizens' profiles. Each individual being classified was assigned to one of these classes according to his answer to the synthetic variable (question 11).

In the next step, the learning LAMDA algorithm was used to learn the citizen profiles from the data concerning the variables 1 to 11.

Finally, the whole process was repeated but adding a new pure qualitative variable and the results were compared.

Three tests were conducted, the first with quantitative treatment of the first 10 variables, the second with a qualitative treatment and the last one with the addition of the qualitative variable obtained from the open answers. Taking as citizens' profiles the four natural classes: very unsatisfied, rather unsatisfied, quite satisfied and very satisfied, the best degrees of correctly classified individuals were only of 53.3 % for quantitative processing, 56.6 % for qualitative treatment and 60 % by incorporating the open answer qualitative information.

A second experiment tried to make classes easier to separate, and consequently their number was reduced to the minimum: satisfied and unsatisfied citizens. Again, the three tests were conducted and the results turn out to be really good this time. Now the global best degrees of correctly classified individuals were of 80 % for quantitative processing, 73.3 % for qualitative treatment and 83.3 % by incorporating the open answer qualitative information.

MinMax, Frank	**Classified as Satisfied**	**Classified as Unsatisfied**
Really Satisfied	**73.3%**	26.7%
Really Unsatisfied	6.7 %	**93.3 %**

Table 1 : Best results with two classes

This means that, from the town council point of view, we can tell that an unsatisfied citizen really is dissatisfied with a 93.3% of accuracy; meanwhile we will success in a 73.3% of the satisfied citizens. Thus, if the Town Council addresses policies to improve satisfaction, only a 6.7 % of unsatisfied people will be unattended.

A more recent experiment on supervised learning has involved the use of RBF neural networks (Domingo et al., 2005). This allows a view of the changes in citizens' satisfaction, analyzing the system's sensibility with respect to variations in input variables. Then, town hall managers would be able to define cost-effective policies in specific directions.

The experiment followed the same steps as the LAMDA experiment. That is, to study and analyze the effect that the use of qualitative information has over RBF performance, two different kinds of training were conducted. Initial training (referred to as quantitative training) only considered the first 10 variables. Second training (qualitative training) adds to the first 10 variables the qualitative variable obtained from the open answers. Moreover, the database was partitioned into six qualitative classes according to the reasons of satisfaction expressed by the citizens' in the survey: Ambient, Aesthetic, Coast & Sea, Nature, Quality of life and Services.

In both training processes, networks learnt on the training set and were tested on the validation set. The results were used to adjust the radial function width (r). To perform this adjustment of the radial width, a total of 4000 simulations were done for each class. Once the radial width was determined, networks were trained on training and validation into six portions, corresponding to the six qualitative sets while the test set is used to assess the generalization ability of the final solution.

Results showed that classification accuracy for the qualitative training (84,4%) was better than for the quantitative training (46.95%). Since the only difference was the use of the qualitative open answers, it seemed to confirm the quality and relevance of this information. Therefore, the use of qualitative values during training and test process not only does not carry a lost of information, but leads to better results, due to the fact that it implies working with the appropriate level of precision.

6 Conclusions and future research

The present work aims at motivating, defining and analyzing the use of different learning algorithms (Artificial Neural Networks, Fuzzy Systems & Hybrid Connectives, Evolutionary Computation and Kernel Methods) in modeling complex systems such as sustainability measuring systems.

Although this paper has focused on a specific indicator of sustainability (Citizen's satisfaction), the methodological aspects considered can be used in a more complex situation.

With regard to open problems and future work, the following comments can be made:

- To define new criteria for representing the information given in input variables
- To compare the results obtained by considering different approaches.
- To apply the given method in multi-classification problems.

References

Aguado J.C. (1998). *A Mixed Qualitative-Quantitative Self-Learning Classification Technique Applied to Situation Assessment in Industrial, Process Control, Doctoral dissertation,* Universitat Politècnica de Catalunya,

Aguado J.C.; et al. (1999). Comparison of structure and capabilities between a non-standard classification technique and the radial basis function neural network. Proceedings of the 13th European Simulation Multiconference (ICQFN 99) Vol. II, 442-448. Warsaw, Poland.

Broomhead D.S.; Lowe D. (1988). "Multivariable Functional Interpolation and Adaptive Networt", *Complex Systems*, vol. 2, pp. 321-355.

Domingo M.; Agell N.; Aguado J.C.; Català A. (2004). *A Qualitative Reasoning Approach for Measuring Citizens'*

Satisfaction. Recent Advances in AI Research and Development. 113, 193-198, IOS Press.

Domingo M.; Agell N.; Parra F.X.; Sánchez M. (2005). *A System for Predicting Citizens' Satisfaction Using Qualitative Information. In Working Notes of the Workshop on Model-Based Systems of the International Joint-Conference on Artificial Intelligence* – IJCAI'05.

European Common Indicators, European Comission, 2006 [Query 06/14/2006] http://europa.eu.int/comm/environment/urban/common_indicators.htm

Ferrer P.; Català A.; Angulo C. (2002). *Estudi dels Indicadors de Sostenibilitat 1 i 6*, ESAII-RR-02-10.

Holland J.H. (1975). "Adaptation in Natural and Artificial Systems", *Univ. of Michigan Press*, Ann Arbor, Michigan.

Hornik K.; Stinchcombe M.; White H. (1989). "Multilayer Feed-forward Neural Networks Are Universal Approximators", *Neural Networks*, 2, pp. 359-366.

Klir J.G.; Yuan B. (1995). *Fuzzy Sets and Fuzzy Logics: Theory and Applications*, Prince Hall PTR.

Poggio T.; Girosi F. (1990). "Networks for Approximation and Learning", *Proceedings IEEE*, vol. 78, pp. 1481-1497.

Vapnik V.N. (1995). *The Nature of Statistical Learning Theory*, Springer-Verlag.

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