A Neuro-fuzzy System for Fraud Detection in Electricity Distribution

C. Muniz, M. Vellasco, R. Tanscheit, K. Figueiredo

Computational Intelligence Lab., Department of Electrical Engineering,
Pontifical Catholic University of Rio de Janeiro
Rio de Janeiro, Brazil
Email: cyroassis@gmail.com, {marley, ricardo, karla}@ele.puc-rio.br

Abstract *— The volume of energy loss that Brazilian electrical utilities have to deal with has been ever increasing. Electricity distribution companies have suffered significant and increasing losses in the last years, due to theft, measurement errors and other irregularities. Therefore there is a great concern to identify the profile of irregular customers, in order to reduce the volume of such losses. This paper presents a combined approach of a neural networks committee and a neuro-fuzzy hierarchical system intended to increase the level of accuracy in the identification of irregularities among low voltage consumers. The data used to test the proposed system are from Light S.A., the distribution company of Rio de Janeiro. The results obtained presented a significant increase in the identification of irregular customers when compared to the current methodology employed by the company.*

Keywords— neural nets, hierarchical neuro-fuzzy systems, binary space partition, electricity distribution, fraud detection.

1 Introduction

Commercial electricity loss due to irregularities (energy theft and measurement errors) reach about R$ 5 billion a year in Brazil, which is around 5% of the total amount spent in energy consumption. A great part of that loss occurs in the area operated by Light S.A.: 3.8 million customers in 28 cities in the state of Rio de Janeiro. Electricity bills are, in some cases, up to 17% more expensive than they should be, so as to compensate for energy thefts and other fraudulent practices [1].

The company makes use of heuristics to indicate a set of low voltage customers who may practice some kind of irregularity. This set of customers, initially classified as suspects, is then analyzed by an expert, who selects a certain number to be directly inspected. The inspection is performed by an employee that visits each residence to confirm (or not) the irregularity. Based on this simple approach, Light S.A. has reached a low identification rate. It is evident that the adopted procedure is not effective and that a more accurate system must be developed.

Some researches have already tackled this problem, confirming the difficulty of correctly identifying irregular consumption. In [2], Rough Sets have been used to fraud detection among electrical energy consumers. The results obtained are promising but accuracy is still low (around 20%) due to the presence of noisy information in the company's database. In [3] and [4] a similar approach is presented, where only past consumption information is considered in the analysis. Due to the temporal nature of the data, time series analysis methods [5] are applied in order to get new invariant characteristics. However, other relevant variables, such as information about the measurement device and the average temperature are not considered.

This paper describes a new intelligent methodology, based on neural networks and fuzzy systems, that takes into consideration historical consumption as well as other important variables. The main objective is to increase the accuracy level when identifying irregular low voltage consumers, selected from a group of suspicious clients.

The proposed methodology is composed of two basic modules: Filtering and Classification, the former comprising an ensemble of five artificial neural networks, while the latter consists of a Neuro-fuzzy Hierarchical System [6]. In the Filtering module, each neural network indicates whether customers are suspected of irregularities or not. As in any ensemble-based methodology, the use of a committee of neural networks aims at improving accuracy by producing a consensus decision that is potentially more accurate than individual neural networks [7]. The Filtering Module was developed to better extract actual normal and irregular customers for training the Classification Module. This module is introduced in the expectation of improving even further the detection of irregularities. It should be noted that the database contains untrustworthy data, due to incorrect information filled in by some visiting employees.

The paper is organized in four additional sections: Section 2 describes the current methodology used by Light S.A. for detecting irregularities, Section 3 describes the proposed methodology, Section 4 presents the case studies considered, and, finally, Section 5 concludes the work.

2 Current Methodology for Fraud Detection

Currently, Light S.A. makes use of diverse methodologies for selecting low voltage customers that present irregularity indications. The *Quarterly Indicator* methodology consists of comparing the past three months' consumption. The *Annual Indicator* consists of evaluating clients' consumption during the last 24 months. The *Adjustment Factor* methodology compares last month consumption to that in the corresponding month of the preceding year. Finally, the *Tendency Factor* methodology consists of comparing clients' previous month's consumption with that expected in the present month. In all cases, if the reduction in consumption is higher than expected, the client is considered a suspect.

All customers identified by those methodologies form

ISBN: 978-989-95079-6-8

the set of suspects, which is analyzed by an expert; some are then recommended for direct inspection. Through this methodology, Light has attained an average *Positive Predictive Value* (PPV) of 25%, which is the proportion of proven irregular customers among all those who have been classified as suspects of some irregularity and have been inspected. Table 1 shows a typical confusion matrix [8] for a two-class problem (*normal* and *irregular* consumers).

Table 1: Suspected x Inspected Customers

System Output	Inspection Result	
	Irregular	Normal
Irregular	A	B
Normal	C	D

The PPV is computed by:

$$PPV = A \big/ A + B$$

3 Proposed Methodology

The methodology proposed in this work consists of three basic modules: Pre-Processing and Normalization, Filtering and Classification.

The whole methodology is shown in Fig. 1. Fig. 1a shows the system's configuration used during the learning phase, which contains all three modules; Fig. 1b, on the other hand, presents a simpler configuration, used for the recall phase, which does not include the Filtering Module.

The first module, present in both training and recall phases and called Pre-processing and Normalization, is where all attributes are selected, normalized and coded.
It should be pointed out, however, that the available inspection database is, unfortunately, very noisy, containing unreliable information due to (intentional or not) incorrect indication of an irregular consumer as a normal one during inspection. Therefore, to create a more reliable database for training the Classification Module, a Filtering Module was conceived.

The Filtering Module makes use of a committee of five neural networks, which select *irregular* and *normal* consumers that better characterize the two different classes.
The resulting filtered database is then employed for training the Classification Module, which eventually identifies a customer as *normal* or *irregular*. Therefore, future classifications of customers (recall phase) will employ only the Pre-processing and Classification modules. The Filtering Module is only necessary if the system needs to be retrained.

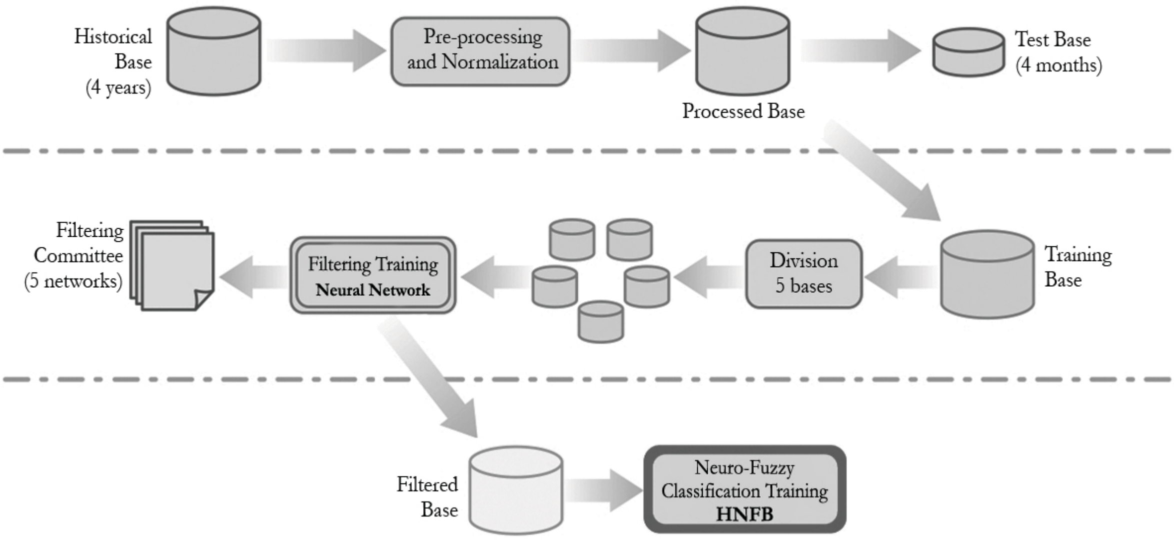

Figure 1(a): Complete methodology of the Learning Phase – Pre-processing, Filtering and Classification

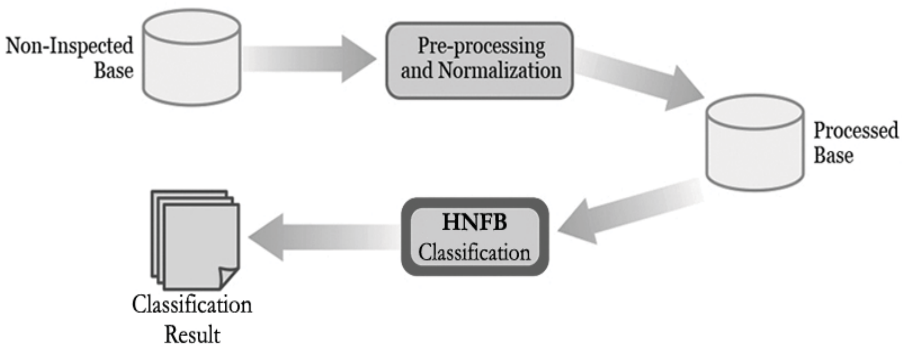

Figure 1(b): Methodology for the Recall Phase - Pre-processing and Classification modules.

ISBN: 978-989-95079-6-8

The following sub-sections present in more detail each of the modules.

3.1 Pre-Processing and Normalization Module

The data pre-processing module is composed of the following phases: selection of attributes, data cleaning and normalization. The database contains general information on both commercial/industrial and residential consumers in the so-called East Regional of Rio de Janeiro. It is well known that the behaviours of residential and non-residential clients are very distinct. Therefore, the database was divided into residential and non-residential clients, resulting in a specific Module for each class of clients.

After the evaluation of the entire available attributes, those which refer to customers' consumption for the last 4 years were selected, as well as some attributes that characterize the installation made at the consumer unit. Table 2 presents the attributes selected in this phase.

Table 2: Attributes selected from the database

Attribute	Description
Local	Consumer Unity Identification code
Inspection date	Date when the inspection was made
Related Month	Month used as a reference. Previous month to the inspection.
Origin	Main reason for the inspection
Measurer	Measuring device identifier
MonoBiTri	Indicates if the installation is mono, bi or triphasic
Consumption irregularity code	Last detected consumption irregularity
Reading irregularity code	Last confirmed reading irregularity
Month consumption	Consumer unity consumption during the related month
Previous year consumption	Registered consumption one year before the related month
Quarterly Indicator	Variation in the past three months' consumption, one by one
Adjustment Factor	Variation from the current month consumption to the same month from the previous year
Annual Indicator	Clients' consumption during the last 24 months
Tendency Factor	Clients' previous month's consumption with the consumption expected in the present month
Fraud indicator	Indicates whether the inspector found an irregularity or not

The first three attributes in the database identify, respectively, the installation where the inspection was made, the date of inspection and the month used as a reference for data acquisition. Those three attributes are presented only to identify consumer registers and are not used as inputs to the system. The five following attributes describe in detail the technical aspects of the installation and the customer's behavior. These are nominal attributes, coded as follows:

Origin – 60 different origins, divided into 4 classes (encoded in 2 binary inputs):

1. Intelligence: information extracted from the database
2. System: automatically identified information by the system of Light S.A
3. Report: complaints made by other customers or readers
4. Grouping: inspection without fraud suspicion

Measuring device – 90 models of measuring devices, grouped in 6 classes (encoded in 3 binary inputs):

1. MMD – direct mechanical measurer
2. OMMD – direct mechanical measurer (obsolete)
3. MMI – indirect mechanical measurer
4. OMMI – indirect mechanical measurer (obsolete)
5. MED – direct electronic measurer
6. OMED – direct electronic measurer (obsolete)

MonoBiTri – three different classes: monophasic, biphasic and triphasic (encoded by 2 binary inputs).

Reading Irregularity Code and *Consumption Irregularity Code* – coded in such a way that if one customer presents any kind of irregularity, his attribute will be 1; otherwise, it is zero.

The next six attributes – *consumption in a given month*, *previous year consumption* and the indicators explained below – define the consumption profile of the customer. Since these are numerical attributes, they have been normalized through:

$$value_{NORM} = \frac{v_{MAX} - value}{v_{MAX} - v_{MIN}}$$

where *value* is the attribute to be normalized, $value_{NORM}$ is the normalized attribute, v_{MAX} is the maximum attribute value registered in the last two years and v_{MIN} is the minimum attribute value registered in the last 2 years.

Four additional indicators have been computed from the monthly consumption information of each customer: *three-months moving average*, *six-months moving average*, *last 12 months average* and the *previous year average* (from the 24th previous month to the 12th previous month).

The *Fraud Indicator* attribute provides the result of the inspection, and indicates if any irregularity was found. This attribute was used as output (target) in training.

Finally, the following variables were also added to the set of attributes: *minimum* and *maximum monthly temperatures*, according to the customer's geographic area, and the *amount of consumption of all customers in the month* under analysis.

All data have been filtered (spurious data, redundant and incomplete were eliminated) and normalized [9]. Through normalization all attributes present the same interval of variation, which is important for training.

ISBN: 978-989-95079-6-8

3.2 Filtering Module

As already mentioned, the available database is rather noisy, due to some unreliable information in the Fraud Indicator attribute. In order to extract patterns from the database that are more reliable examples of actual *normal* and *irregular* customers, the Filtering Module was proposed. This consists of a committee of five neural network, all of them Multi-Layer Perceptrons (MLP) [10][11], with 21 inputs, one hidden layer and one neuron in the output layer, with sigmoidal threshold function (Fig. 2). The output neuron represents two classes: *irregular* customers (measurement or technical irregularities) and *normal* customers (those who do not show any irregularity).

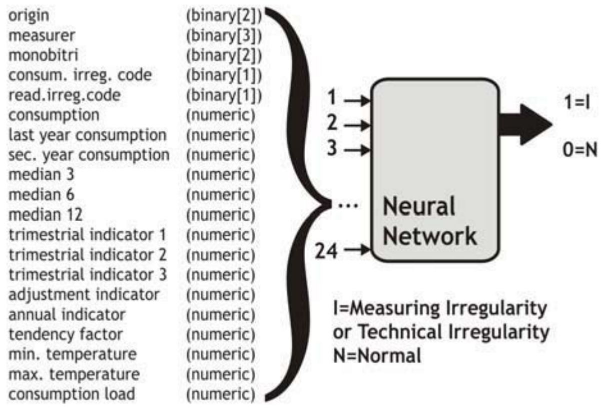

Figure 2: MLP Neural Network topology with one output

As the network output is a continuous value between zero and one, a threshold level of 0.5 was defined so as to distinguish *irregular* (output value ≥ 0.5) from *normal* customers (output value < 0.5). It should be mentioned that different threshold limits, between [0.3, 0.8] with steps of 0.05 were also evaluated, but the best performance has been achieved for 0.5.

The complete database is divided into five disjoint sets in order to train each of the networks that form the committee. After training, the complete database is processed by the neural networks. The selection of more reliable training patterns (filtering process) is carried out as follows: a pattern that is labeled as *irregular* in the dataset and is indicated by three or more votes in the committee evaluation as also *irregular* is taken as a positive true pattern, that is, a customer that actually characterizes an *irregular* pattern. On the other hand, patterns specified as *normal* in the database and indicated by at least three votes from the committee as being *normal* are called negative true patterns. These patterns are then considered as reliable examples of normal customers (see Fig. 3).

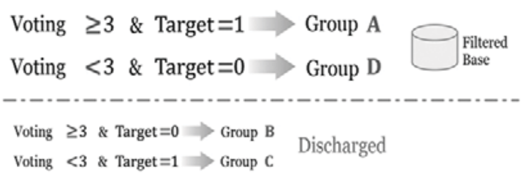

Figure 3: Limits for the Filtering Committee

A new filtered database is then formed, containing all positive and negative true patterns, which will then be used for training the Classification Module. The positively false registers (those normal patterns that receive three or more votes as being irregular from the committee) and the negatively false registers (irregular patterns that receive at least three votes as being normal) are discarded as noisy patterns and are not used to train the Classification Module.

3.3 Classification Module

As mentioned above, the main difference between this and the filtering module is the database used for training. Now only the filtered database, obtained from the filtering module, is used to adjust the classifier parameters. Once training is completed, the Classification Module is ready for use with new input patterns. In this phase, the Filtering Module is no longer necessary (see Fig. 1b), and will only be used again if retraining becomes necessary.

The structure used in this stage is a Hierarchical Neuro-fuzzy with Binary Space Partition (BSP) system [6] [12], briefly described below.

Hierarchical Neuro-Fuzzy Systems have been devised to overcome limitations of traditional neuro-fuzzy systems, which in general have a limited capacity for creating their own structure and rules. Additionally, most of the models employ grid partition of the input space, which, due to the rule explosion problem, are more adequate for applications with a smaller number of inputs. When a greater number of input variables are necessary, the system's performance deteriorates.

The Neuro-fuzzy Hierarchical BSP (HNFB) system used here makes use of basic cells. An HNFB cell is a neuro-fuzzy mini-system that performs fuzzy binary partitioning of the input space. The HNFB cell generates a crisp output after a defuzzification process. Fig. 4 illustrates the cell's functionality, where x represents the input variable, $\rho(x)$ and $\mu(x)$ are the membership functions *low* and *high*, respectively, which generate the antecedents of the two fuzzy rules, and y is the output. The linguistic interpretation of the mapping implemented by the HNFB cell is given by the following rules:

- If $x \in \rho$ then $y = d_1$
- If $x \in \mu$ then $y = d_2$

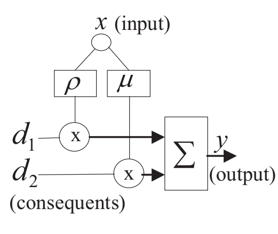

Figure 4: Neuro-Fuzzy BSP cell

Each rule corresponds to one of the two partitions generated by BSP. Each partition can in turn be subdivided into two parts by means of another HNFB

ISBN: 978-989-95079-6-8

cell. The profiles of membership functions $\rho(x)$ and $\mu(x)$ are complementary logistic functions.

The output y of an HNFB cell (defuzzification process) is given by the weighted average. Due to the fact that the membership function $\rho(x)$ is the complement to 1 of $\mu(x)$, the following equation applies:

$$y = \rho(x) * d_1 + \mu(x) * d_2 \quad \text{or} \quad y = \sum_{i=1}^{2} \alpha_i d_i$$

where α_i symbolizes the firing level of the rule in partition i and are given by $\alpha_1 = \rho(x)$ and $\alpha_2 = \mu(x)$. Each d_i corresponds to one of the possible consequents below:

- A singleton: d_i = constant.
- A linear combination of inputs: $d_i = \sum_{k=1}^{n} w_k x_k + w_0$

  where: x_k is the k-th input, w_k represent the weight associated with input x_k, n is the total number of inputs and w_0 is a constant value.
- The output of a generic cell j: $d_i = y_j$

The HNFB model may be described as a system that is made up of interconnections of HNFB cells, as illustrated in Fig. 5, along with the respective partitioning of the input space. In the system presented in Fig. 5, the initial partitions 1 and 2 ('BSP0' cell) have been subdivided; hence, the consequents of its rules are the outputs of BSP1 and BSP2, respectively. In turn, these subsystems have, as consequents, values d_{11}, y_{12}, d_{21} and d_{22}, respectively. Consequent y_{12} is the output of the 'BSP12' cell. The output of the system in Fig. 3(b) is:

$$y = \alpha_1(\alpha_{11} d_{11} + \alpha_{12}(\alpha_{121} d_{121} + \alpha_{122} d_{122})) + \alpha_2(\alpha_{21} d_{21} + \alpha_{22} d_{22})$$

Although each BSP cell divides the input space only into two fuzzy sets (*low* and *high*), the complete HNFB architecture divides the universe of each variable into as many partitions as necessary. The number of partitions is determined during the learning process. In Fig. 5, for instance, the upper left part of the input space (partition 12) has been further subdivided by the horizontal variable x_1, resulting in three fuzzy sets for the complete universe of discourse for this specific variable.

The training algorithm makes use of the gradient descent method for learning the structure of the model and, consequently, linguistic rules. The parameters that define the the membership functions of the antecedents and consequents are regarded as fuzzy weights of the neuro-fuzzy system.

A tuning parameter δ, called decomposition rate prevents the structure from growing indefinitely.

4 Experimental Results

This section presents the results obtained for Filtering, the Neural Networks Committee, and Classification, performed by the HNFB system.

4.1 Filtering

After processing the selected data three databases were obtained. These were composed of 4159 customers with non technical irregularity, 3754 customers with technical irregularity and 14405 normal customers.

Data were divided into Training, Validation and Testing databases. Training and Validation sets refer to the period between September 2002 and June 2006, with the exception of February and March 2006. The data for these two months, together with those for July and August 2006 were used to form the test set. These months were selected to evaluate the system's performance in different seasons (summer and winter in the southern hemisphere). Due to the unbalanced characteristic of the database, with many more examples of *normal* customers than the available number of *irregular* patterns (technical and non-technical irregularities), the training and validation sets were equalized to avoid a biased behaviour of the neural networks. The datasets actually used in the experiments consisted of 3754 customers with technical irregularities, 3754 randomly selected customers with non-technical irregularities and 7508 normal ones, also randomly chosen. For training, 75% of the balanced samples have been considered; the rest was left for validation.

Five different databases were created, each to train one of the five neural networks that form the filtering committee. For each neural network, 10 training processes were realized and the best performing network from the validation set has been selected to form the committee. The early stopping methodology was used during each training process, which specifies the optimal number of epochs to avoid over fitting and attain best generalization performance.

As the main objective of the proposed system was to increase the PPV of irregularity detection, two experiments were carried out: one considering the PPV as the error metric for the neural networks training, and another considering the general classification error as the metric used to select the best neural network. It was observed that the latter provided better results, since when the classification error for both classes is minimized, the PPV is indirectly maximized. Thus, in the following experimental results, the general classification error metric was used.

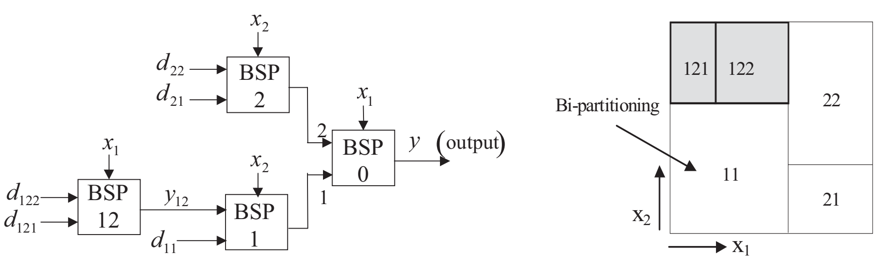

Figure 5: Example of an HNFB system and respective input space Partitioning

ISBN: 978-989-95079-6-8

After all neural networks had been trained and the committee members had been selected, a test was carried out with the data of customers investigated during the months of February, March, July and August 2006. That test base has 4663 registers of inspected customers.
Table 3 illustrates filtering results for one specific ZEI (Elementary Intervention Zones) of East Regional, in the same fashion of Table 1.

Table 3: Results with the filtering committee

	Committee	
	Irregular	Normal
Irregular	960	1164
Normal	572	1970
	PPV_{NN}:	45.2%
	Class. Error:	37.2%
	PPV_{Light}:	32.8%

As can be observed, by using only the Filtering Module the PPV has already increased from 32.8% to 45.2%.

4.2 Classification

The same inputs considered for filtering were initially considered for classification, performed by the HNFB system. After variable selection by the least-square estimator (LSE) [13], nine were effectively used as inputs to the system. The test database was identical to that used for testing in the filtering stage.

It can be seen from the classification results shown in Table 4 that the VPP has improved even further, reaching 51.2%. This attests the advantage of the proposed methodology.

Table 4: Results for classification

	Committee	
	Irregular	Normal
Irregular	991	943
Normal	541	2191
	PPV_{HNFB}:	51.2%
	Class. Error:	31.8%
	PPV_{Light}:	32.8%

In addition to performing classification, the HNFB also produces fuzzy rules as a result. In this experiment 45 rules have been generated, of which an example is:
If *Origin* is *low* and *minimum temperature* is *low* then class is *normal*.

5 Conclusions

This paper presented an intelligent system, based on a neural networks committee and on a Hierarchical Neuro-fuzzy BSP system, to identify consumers' frauds in an electrical distribution company of Rio de Janeiro, Brazil.

The system is formed by two modules, one for filtering the database and another for actual classification of a consumer. The use of this methodology greatly improved the current fraud detection rate.

Besides increasing the performance regarding the correct identification of fraudulent consumers, the system has the additional advantage of ranking irregular costumers as a result of the number of votes the neural networks committee provides to identify a customer as irregular.

The proposed methodology is being applied to other areas covered by the company. Additionally, a clustering algorithm will be applied to non-residential consumers in order to group them in classes with similar consumption behaviour.

References

[1] J. E. C. Junior, *Detecção de Fraudes em Unidades Consumidoras de Energia Elétrica Utilizando Rough Sets*, M.SC. Dissertation, UFMS (in Portuguese), 2005.

[2] J. E. Cabral, E. M. Gontijo, J. O. P. Pinto and J.R. Filho, Fraud Detection in Electrical Energy Consumers using Rough Sets, *IEEE International Conference on Systems, Man and Cybernetics*, 4: 3625-3629, 2004.

[3] T. Rauber, I. Drago, F. Varejão and R. Queiroga, Extracão e Seleção de Características na Identificação de Perdas Comerciais na Distribuição de Energia Elétrica, *XXV Congress of the Brazilian Computation Society* (in Portuguese), 2005.

[4] J. Rong, H. Tagaris, A. Lachsz and M. Jeffrey, Wavelet based Feature Extraction and Multiple Classifiers for Electricity Fraud Detection, *Transmission and Distribution Conference and Exhibition*, 3: 2251-2256, 2002.

[5] D. S. G. Pollock, *A Handbook of Time-Series Analysis, Signal Processing and Dynamics*. Academic Press, New York, 1999.

[6] M. Vellasco, Pacheco, K. Figueiredo, F. Souza, Hierarchical Neuro-Fuzzy Systems - Part I, *Encyclopedia of Artificial Intelligence*, Juan R. Rabuñal, Julián Dorado & Alejandro Pazos (Eds.), Information Science Reference, 2008.

[7] R. Santos, M. Vellasco, F. Artola and S. Fontoura. Neural Net Ensembles for Lithology Recognition, *4th International Workshop on Multiple Classifier Systems*, LCNS 2709 – Multiple Classifier Systems, Springer-Verlag, 2003.

[8] F. Provost, T. Fawcett, and R. Kohavi, (1998), The Case Against Accuracy Estimation for Comparing Induction Algorithms, *15th International Conference on Machine Learning*, 445-453, 1998.

[9] M. Dash and H. Liu, Feature Selection for Classification, *Intelligent Data Analysis*, Elsevier, 1(3): 131 - 156, 1997.

[10] S. Haykin (1999), *Neural Networks - a Comprehensive Foundation*, Prentice-Hall, New Jersey, 1999.

[11] C.M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, 1995.

[12] L.B. Gonçalves, M.M.B.R. Vellasco, M.A.C. Pacheco, F. J. de Souza, Inverted Hierarchical Neuro-Fuzzy BSP System: A Novel Neuro-Fuzzy Model for Pattern Classification and Rule Extraction in Databases, *IEEE Transactions on Systems, Man & Cybernetics, Part C: Applications and Review*, 36(2): 236-248, March 2006.

[13] F.-L. Chung and J.-C. Duan, On Multistage Fuzzy Neural Network Modeling, *IEEE Transactions on Fuzzy Systems*, 8(2): 125-142, 2000.

ISBN: 978-989-95079-6-8