Artificial Neural Networks, Multiple Linear Regression and Decision Trees Applied to Labor Justice

Genival Pavanelli, Maria Teresinha Arns Steiner,
Alessandra Memari Pavanelli and Deise Maria Bertholdi Costa
Specialization Program in Numerical Methods in Engennering PPGMNE,
Federal University of Paraná State UFPR, CP 1908 , Curitiba, Brazil

Keywords: Mathematical Programming, Artificial Neural Networks, Multiple Linear Regression, Decision Tree, Principal Component Analysis, Encoding Attributes.

Abstract: This paper aims to predict the duration of lawsuits for labor users of the justice system. Thus, we intend to provide forecasts of the duration of a labor lawsuit that gives subsidies to establish an agreement between the parties involved in the processes. The proposed methodology consists in applying and comparing three techniques of the Mathematical Programming area, Artificial Neural Networks (ANN), Multiple Linear Regression (MLR) and Decision Trees in order to obtain the best possible performance for the forecast. Therefore, we used the data from the Labor Forum of São José dos Pinhais, Paraná, Brazil, to do the training of various ANNs, the MLR and the Decision Tree. In several simulations, the techniques were used directly and in others, the Principal Component Analysis (PCA) and / or the coding of attributes were performed before their use in order to further improve their performance. Thus, taking up new data (processes) for which it is necessary to predict the duration of the lawsuit, it will be possible to make up conditions to "diagnose" its length preliminarily at its course. The three techniques used were effective, showing results consistent with an acceptable margin of error.

1 INTRODUCTION

This work presents a proposal of application of techniques in the field of Operational Research, by the labor courts. This proposal is to provide an estimate of the duration of a labor lawsuit for users of the Labor Forum of Sao Jose do Pinhais, PR, Brazil.

In order to obtain such a prediction, we used three methods: one from the area of artificial intelligence, Artificial Neural Networks (ANN) and two, from the Statistical Area, Multiple Linear Regression (MLR) and Decision Tree. The purpose of using these three methods already well known among search sources is to make a comparison between the final results and, thus, determine which provides the best performance (highest percentage of correct answers) and thus be used in future forecasts.

This paper is structured as follows: section 2 presents related work that also made use of Operational Research techniques applied here. Section 3 is a description of the problem, gathering and processing of data. Section 4 presents the methodology of the work, which describe the concepts involving the techniques of ANNs, Principal Component Analysis (PCA), MLR and Decision Tree. Section 5 describes the implementation of computational techniques and analysis of results. Finally, section 6 presents the conclusions obtained by analyzing the results of the previous section.

2 RELATED WORK

There are in literature, many studies related to data forecasting, in which various techniques in the field of Operations Research and, more specifically, Pattern Recognition, have been applied. It is noteworthy that no studies were found related to forecasting problems of the Labor Court, as presented here. Among the studies reviewed in the literature, may be mentioned those listed below.

In Baptistella, Cunico and Steiner (2009), the

Pavanelli G., Teresinha Arns Steiner M., Memari Pavanelli A. and Maria Bertholdi Costa D..
Artificial Neural Networks, Multiple Linear Regression and Decision Trees Applied to Labor Justice.
DOI: 10.5220/0004517504430450
In *Proceedings of the 5th International Joint Conference on Computational Intelligence* (NCTA-2013), pages 443-450
ISBN: 978-989-8565-77-8

authors look for alternative techniques in order to determine market values for properties in Guarapuava, PR. It is proposed the use of ANNs and for that, we collected 256 historical records (patterns) of urban real estate in the city. Each of the records was composed of 13 information (attributes): neighborhood, sector, paving, drainage, street lighting, land area, soil conditions, topography, location, built up area, type, structure and conservation. Several simulations have been developed, with the worst results presenting an accuracy of 78% and the best, 95%.

Still in property valuation, there is also the work of Nguyen and Cripps (2001), which compares the performance of ANNs with Multiple Regression Analysis for the sale of family houses. Multiple comparisons were made between the two models in which were varied: the sample size of data, the functional specification and the time prediction. In the work of Bond; Seiler and Seiler (2002), the authors examine the effect that the view of a lake (Lake Erie, USA) has on the value of a house. In the study the transaction prices of houses were taken into account (market price). The results indicate that, in addition to the variable view, which is significantly more important than the others, also the building area and the batch size are important.

Baesens et al. (2003) discuss three methods for the extraction of rules from a neural network, in a compared way: NeuroRule; Trepan and Nefclass. To compare the performances of the methods discussed, we used three real credit datasets: German Credit (obtained from UCI repository), Bene 1 and Bene 2 (obtained from the two largest financial institutions in the Benelux). The algorithms mentioned are also compared with C4.5-tree algorithms, C4.5-rules and Logistic Regression. The authors also show how the extracted rules can be viewed as a decision table in the form of a compact and intuitive graph, allowing better reading and interpretation of results to credit manager.

In Mota and Steiner (2007), the authors present a methodology composed of Multivariate Analysis techniques, to build a statistical model of Multiple Linear Regression for property valuation. It is applied, initially, the Cluster Analysis to the data of each class of urban real estate (apartments, houses and land) to obtain homogeneous groups within each class, and in correspondence, are determined discrimnants to allocate future items in these groups, by the Quadratic discriminant Score Method. Then, it is applied the PCA technique to solve the problem of multicollinearity that may exist between the variables of the model. With the scores of the principal components it is adjusted a Multiple Linear Regression model for each group of homogeneous properties within each class. The methodology was applied to a set of 119 buildings (44 apartments, 51 houses and 24 lots), the city of Campo Mourão, PR. The model for each homogeneous group within each class of property assessed had a proper fit to the data and a predictive quite satisfactory.

Adamowicz (2000) uses pattern recognition techniques, ANN and Linear Discriminant Analysis of Fisher, with the goal of classifying companies as solvent or insolvent. The data were provided by the Southern Regional Development Bank (BRDE), Curitiba branch, PR. Both techniques were efficient in discriminating the companies, and the performance of ANNs was slightly better than the Linear Discriminant Analysis of Fisher.

Ambrosio (2002) presents a study that aims to develop a computer system to assist radiologists in the confirmation of diagnosis of interstitial lung lesions. The data were obtained from the Hospital das Clínicas of the Medicine University of Ribeirão Preto (HCFMRP) using protocols generated by experts. The system was developed using multilayer ANNs as a pattern classifier. The training algorithm is back-propagation with the sigmoidal activation function. Several tests were performed for different network configurations. It was clear that the use of this tool is feasible, since once the network is trained and the weights set, it is no longer necessary to access the database. This makes the system faster and computationally lighter. The research concludes that the ANNs fulfill well their role as classifiers standards.

Souza et al. (2003) used ANNs techniques with three layers of neurons with the back-propagation algorithm. The goal was to predict the content of mechanically separated meat (CMS) in meat products from the mineral content contained in sausages formulated with different levels of chicken. The technique proved to be very efficient during the training and testing, however, the application of the ANNs to commercial samples was inadequate, because of the difference in the ingredients used in the sausage of the training and the ingredients of the commercial samples.

In Steiner, Carnieri and Stange (2009), it is proposed the use of a Linear Programming Model for Pattern Recognition of paper reels of good or poor quality. Data were collected from 145 rolls of paper (standard), 40 of good quality and 105 of low quality. From each coil 18 attributes were considered: tensile and tear tests of pulp, mechanical pulp and thermo-mechanical pulp; amounts of these three folders; consistency and flow of pulp and seven data press rolls of the paper machine. From the PL model, it was built a second mathematical model that makes use of the first, so as to ensure the attainment of good quality coils at a minimum cost.

Biondi Neto et al. (2006) show in their work that the determination of soil type, until then, could be obtained using abacuses; the aim of the research was to apply a computational method to classify the soil. The technique used was again ANNs with Levenberg-Marquardt method, which has resulted in the classification of soil for each increment of depth. All data were obtained from real situations. The convergence time was quick, which facilitated the completion of several tests.

Lu; Setiono and Liu (1995, 1996) reported in their articles the algorithm called Neurorule that makes the extraction of rules from a trained neural network, obtaining rules of the type IF-THEN. The performance of this approach is verified in both articles in an issue of bank credit, and to facilitate the extraction of such rules, the values of numeric attributes were discretized by dividing them into sub-intervals. After the discretization, the encoding scheme "thermometer" was employed to obtain binary representations of the intervals previously defined obtaining thereby the inputs to the neural network. The results obtained indicate that, using the proposed approach, high quality rules can be discovered from a data set.

Steiner et al (2007) use rules extraction techniques such as Neurorule and WEKA software to extract rules from a trained artificial neural network. The ANN classifies companies as good or bad credit borrowing. From the trained network the authors conducted three types of tests for extracting rules. In the first, the extraction of rules was made directly from the original data, in the second test patterns misclassified by the RN were discarded, while in the third test, in addition to discarding patterns misclassified by RN, attributes were coded according to the encodings "thermometer" and "dummy", making them binaries. The results were quite satisfactory presenting accuracy above 80% for the grant (or not) of bank credit.

And so, several other studies from different research areas, making use of various techniques of Pattern Recognition, especially ANNs, could be cited here.

3 DESCRIPTION OF THE PROBLEM

Currently, many countries have labor laws, but was not always so. In Brazil, labor courts and labor law emerged only after the nineteenth century, after many struggles and demands from the working classes. Only after the Revolution of 1930 the Ministry of Labor was created, and the Labor Court was provided by the 1934 Constitution. Currently the Labor Court is structured in three levels of jurisdiction:

- First Level: Labor Courts;
- Second Level: Regional Labor Courts;
- Third Level: Superior Labor Court.

According to the Superior Labor Court (SLC), in Brazil there are 24 Regional Labor Courts (RLC), and as of 2003, about 270 new Labor Courts were created in order to accelerate the legal procedures of labor lawsuits (SLC, 2007). Only in the state of Paraná, at the 9th Region of the RLC, there are 28 Justices distributed statewide (TRT, 2007). Of the 77 Labor Courts of the State of Paraná, São José dos Pinhais (SJP) ranks second in number of labor lawsuits. In 2006, the SJP Forum of Labor started having a second Labor Court. Due to the increasing number of labor lawsuits as a result of massive industrialization in the municipality, it is necessary agility in service of justice. Thus, the use of mathematical tools for predicting the duration labor lawsuits is of fundamental importance to this optimization of time.

The process data (patterns), as well as the attributes of each pattern, used for the development of this work were obtained from the First Labor Court Board of SJP, PR, Brazil. Aiming to determine which attributes would be relevant in determining the duration of a labor lawsuit, several meetings were held with the titular judge of this Forum. As a result of these meetings, we came to a set of 10 attributes listed below.

a. Rite: which may be of labor (LR) or a summary lawsuit (SL)

b. Service time: is the difference between the date of admission and date of discharge, in months;

c. Salary of the Complainant: last salary received;

d. Profession: function performed by the complainant. This attribute was divided into two parts: a sector that is also divided into commerce, industry and service; and office position, which falls into the direction and execution;

e. Process Goal: corresponds to the requests made by the complainant. They can be: lack of registration with professional portfolio, wage differentials, severance, Art 477 fine, Art 467 fine, overtime and reflexes, guarantee fund for length of service, compensation for moral damages, unemployment insurance, transportation payment, health hazard allowance, night allowance and health plan;

f. Agreement: when there is an agreement between the parties;

g. Expertise: whether or not there is a need of performing some kind of expertise, for example, a medical examination or health hazard examination;

h. Regular feature: when one party (plaintiff or claimed) does not agree with the sentence issued by the judge and asks ordinary appeal to the SLC;

i. Review feature: when one party (plaintiff or claimed) does not agree with the judgment of the SLC and requires the Review feature;

j. Number of Hearings: refers to the number of hearings necessary for the judge to issue the sentence;

The 10 attributes listed above, used to predict the duration of the process were collected from 100 cases generating the matrix intended for training and testing of ANNs, as well as for applying the technique of MLR and the construction of the Decision Tree.

Most data was treated to correspond to one or more binary coordinates (Lu; Setiono and Liu, 1996), (Baesens et al., 2003) of the inputs vector to the techniques used, as mentioned in section 3.1, the below.

3.1 Encoding of Attributes

In order to try to improve the performance of techniques, each of the 10 attributes cited was "treated" so as to correspond to one or more binary coordinates (Lu et al., 1996), (Baesens et al., 2003) depending on whether it was nominal or ordinal. We used a "thermometer coding" for the ordinal attributes and "dummy coding" (artificial) for nominal attributes (Baesens et al., 2003), (Steiner et al., 2007).

Table 3.1 illustrates the "thermometer encoding" for the ordinal attribute "Salary of the Complainant", for example. This attribute is first discretized in the values of 1 to 5; for example, the "Input 1 = 1", this means that the original variable "Salary of the Complainant "> 1340. Table 3.2 illustrates the "dummy coding" for the nominal variable "Agreement", for example.

Table 3.1: An example of "thermometer encoding" for ordinal variables.

"Salary of the Complainant" SR(reais)	Cate goric Input	Input 1	Input 2	Input 3	Input 4
$330 \leq SR < 450$	1	0	0	0	0
$450 \leq SR < 620$	2	0	0	0	1
$620 \leq SR < 800$	3	0	0	1	1
$800 \leq SR \leq 1.340$	4	0	1	1	1
$SR \geq 1.340$	5	1	1	1	1

Table 3.2: An example of "dummy coding" for ordinal variables.

Original Input "Agreement"	Input 1
Agreement = Yes	0
Agreement = No	1

From the above-explained encoding the 10 attributes provide 32 inputs to the ANN, therefore, the matrix has 100 rows and 32 columns.

4 METHODOLOGY OF WORK

The methodology applied in this study sought, through the use of ANNs, the MLR and Decision Tree, comparatively recognize patterns in labor lawsuits analyzed to predict the length of the labor lawsuits users of the justice system, as already mentioned.

Aiming to minimize the error of the techniques applied, three different tests were carried out. In the first test all attributes were coded as described in section 3.1, so that each pattern would present an input vector with 32 binary coordinates. In the second test the coded data matrix (according to the previous test) was submitted to PCA, in order to evaluate the relative importance of the variables in the sample data. In the third test the original ordinal variables were not coded, in other words, the attributes salary, service time and number of audience have not been converted into binary vectors and then the matrix was subjected to the PCA, such that each pattern would present, for this test, an input vector of 23 coordinates.

4.1 Artificial Neural Networks

The ANN implemented in this work, classified as multiple layers network or feed-forward network, was trained by back-propagation algorithm using the sigmoidal transfer function, which generates output between "0" and "1" for inputs between $-\infty$ e $+\infty$, in all neurons. Network performance was verified through the MSE (medium quadratic error), given by equation (4.1).

$$MSE = \frac{\sum_{i=1}^{n}\left(d_i^P - a_i^P\right)^2}{2n} \tag{4.1}$$

where n = number of patterns, d_i is the desired output (real value) for the default p e a_i is the output obtained by the network, for the default p.

4.2 Multiple Linear Regression

The second method used in this work, MLR has as main objective to describe the relationship between a response variable and one or more explanatory variables. The most commonly used types of regression are: Linear and Logistic, widely used in various fields of knowledge.

According to Lima (2002), in 1845 the logistic regression technique arose in order to solve problems of population growth. This technique has also been employed in the field of Biology in the 30s. However, its application in economic and social problems appears only in the 60s. Recently, this methodology has become mandatory in many reference econometrics manuals. Logistic Regression is a statistical technique widely used in data analysis with responses belonging to the interval [0, 1], with the goal of classifying patterns into classes.

Linear Regression is widely used in many areas of research, being a kind of technique that can produce values of estimated response outside the range [0, 1]. It is considered a classical regression model. It is a technique used to study the relationship between one dependent variable and several independent variables. The goal can be explanatory, i.e., demonstrate a mathematical relationship that can indicate but not prove a cause and effect relationship, or predictive, i.e., obtain a relation that permits, through future observations of the variables x, predict the corresponding value y.

Suppose you want to build a model that relates the response variable y with p factors $x_1, x_2,...,x_p$. This model always includes an error range. There is then:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_1 x_{i2} + ... + \beta_p x_{ip} + \varepsilon$$

for $i = 1,2,..,n$ where n is the number of observations; p is the number of variables.

Using matrix notation: $Y = X\beta + \varepsilon$, where Y is the response variable, X matrix model; β is the vector of parameters to be estimated; ε is the vector of random errors.

$$Y = \begin{bmatrix} \overline{Y}_1 \\ \overline{Y}_2 \\ \vdots \\ \overline{Y}n \end{bmatrix} \quad X = \begin{pmatrix} 1 & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{np} \end{pmatrix} \quad \beta = \begin{bmatrix} \overline{\beta}_0 \\ \overline{\beta}_1 \\ \vdots \\ \overline{\beta}_p \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \overline{\varepsilon}_1 \\ \overline{\varepsilon}_2 \\ \vdots \\ \overline{\varepsilon}_n \end{bmatrix}$$

4.3 Decision Trees

Decision trees are a very powerful technique, widely used, based on a hierarchy of tests to some of the variables involved in a problem of decision. The knowledge gained from this technique is expressed through rules, a fact that justifies its widespread use. It can be used for two purposes: **prediction** (example: find out if a customer will be a good payer according to his/her characteristics) and **description** (provide interesting information about the relationships between predictive attributes and class attribute in a database).

Its structure has the following characteristics:

- Each internal node is a test on a predictive attribute;
- a branch starting from an internal node represents a result for the test;
- a leaf of the tree represents a class label.

To classify an unknown example it is just necessary to forward it down the tree according to the values of the attributes tested in successive nodes, and when a leaf is reached the instance is classified according to the class assigned to the leaf (Witten and Frank, 2005).

4.4 Principal Component Analysis

Looking for further improvement of the obtained results, in some of the tests it was applied the PCA, which is able to identify patterns in data, in order to express them pointing out the similarities and also the existent differences. It is linked to the covariance structure explanation through a few linear combinations of the original variables.

Furthermore, through PCA the original size of the database is reduced with linear combinations of a set of variables that retain the maximum number of information contained in the original variables, and also facilitates the interpretation of the analyzes, judging the importance of the original variables chosen.

5 COMPUTING IMPLEMENTATION AND ACHIEVEMENT OF RESULTS

As previously described in section 4, the methods proposed in this paper, ANNs, MLR and Decision Trees have been applied after collection and processing (coding of attributes / implementing PCA) of the cases examined, which were filed at the Forum of Labor of SJP. All data obtained in each lawsuit were used to compose the input matrices. The training of the ANNs implemented in this work is supervised, i.e. for each data input vector the output is already known (Haykin, 2002). Thus, in order to perform the training and testing of ANN it was implemented a program using Visual Basic 6.0 Software.

To carry out the training of ANNs it was used, as already mentioned, the back-propagation supervised algorithm, with sigmoidal activation function, with outputs in the range [-1, 1]. Due to these conditions of the activation function, it was necessary to fit the outputs in this interval. Thus, the length of processes which range from 1 to 94 months, were divided by 94. It is noteworthy that the length of the process has a uniform distribution.

For the assessment of the ANN, it was used the hould out procedure, in other words, from all the processes registered, 75% were used for the training of the network and the remaining 25% were used in the test. In the application of ANNs, four sets of initial weights were used in all tests.

Figure 5.1, below, shows conceptual forms of two learning curves, one in relation to the training set and another in relation to the validation set. The curves are distinctive, and the curve of training (learning) monotonously decreases for an increasing number of iterations. Since the validation curve decreases to a minimum and then increases again as the training proceeds.

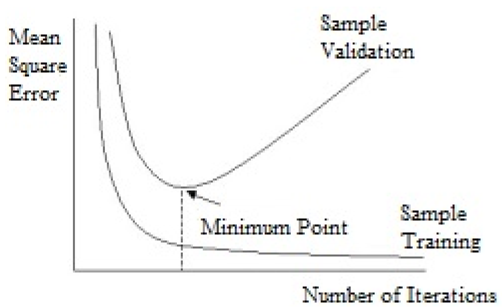


Figure 5.1: Training Versus ANN Generalization Capability - Source: Haykin (2002).

In all tests performed on ANNs, it was varied the number of hidden layer neurons from "1" to "15", remaining fixed the number of 50 iterations for each of the topologies in order to find the lowest error in the test group. The architecture that provided the smallest error, returned to be trained, now varying the number of iterations, until the moment when the error in the test group reached the minimum. Thus, the over-fitting of the network will be avoided, in other words, the ANN would give better results for the training group. However, it would lose the ability to generalize, as illustrated in figure 5.1.

A nomenclature has been chosen for each topology, in order to represent, in sequence, the following characteristics of the ANN: number of entries, the number of neurons in the hidden layer and the number of iterations. For example, the network "E32N3I40" is a network with "32" entries "3" neurons in the hidden layer and was trained with "40" iterations.

5.1 Results obtained

In predicting the length of a labor lawsuit, the third test showed the best result, where ordinal attributes were not coded and then the data were submitted to PCA. Table 5.1 below shows the variation in the number of neurons in ANNs for this test.

Table 5.1: Results of simulations varying the Number of Neurons in Hidden Layer.

SIMU LATION	TOPO-LOGY	MSE Tr	MSE Tes
1	E23N1I50	0,06968	0,14800
2	E23N2I50	0,05303	0,20287
3	E23N3I50	0,03189	0,16561
4	E23N4I50	0,03207	0,13947
5	E23N5I50	0,02686	0,36916
6	**E23N6I50**	**0,02383**	**0,07108**
7	E23N7I50	0,02294	0,08402
8	E23N8I50	0,03090	0,12042
9	E23N9I50	0,02282	0,25118
10	E23N10I50	0,02502	0,45144
11	E23N11I50	0,02375	0,14043
12	E23N12I50	0,02225	0,12329
13	E23N13I50	0,02485	0,11329
14	E23N14I50	0,02186	0,21053

According to table 5.1, it appears that the best network topology is E23N6I50, which means, 23 neurons in the input layer, six in the hidden layer, trained with 50 iterations. From this analysis, this network has been trained, now varying the number of iterations in order to obtain the lowest possible error.

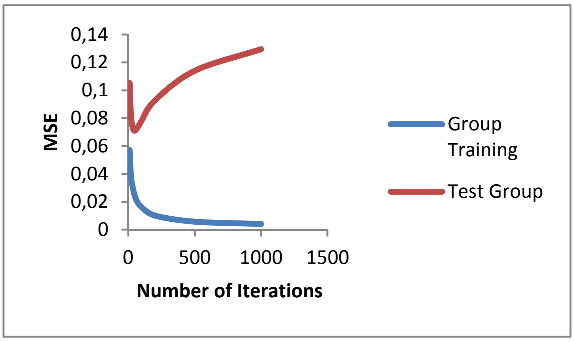
From the results of table 5.2, we can see that the simulation with 50 iterations provides the best results (lower error rate in the test group).

Table 5.2: Results of simulations varying the number of iterations.

SIMU LATION	TOPO-LOGY	MSE Tr	MSE Tes
1	E23N6I10	0,05718	0,10532
2	E23N6I20	0,03792	0,08161
3	E23N6I30	0,03175	0,07455
4	E23N6I40	0,02738	0,07145
5	**E23N6I50**	**0,02383**	**0,07108**
6	E23N6I60	0,02134	0,07196
7	E23N6I70	0,01955	0,07333
8	E23N6I80	0,01819	0,07489
9	E23N6I90	0,01706	0,07653
10	E23N6I100	0,01609	0,07821
11	E23N6I200	0,01014	0,09270
12	E23N6I500	0,00573	0,11402
13	E23N6I1000	0,00409	0,12951

As expected, the error in the training group decreased monotonously at the same time that the number of iterations increases. In the test group the error decreases reaching a minimum of 0.07108 when the network is being trained with 50 iterations. When we increase this number it becomes very clear that the error in this group begins to increase characterizing the loss of generalization capability of ANN from that moment. Such information can be seen in Graph 5.1 below.

Graph 5.1: MSE Training group and test group.



The prediction made through the MLR technique, used the same data sets (training and test) of ANNs, as well as three types of tests in order to compare the results.

The best result obtained with this technique was also the third test. The vector of estimated parameters obtained in the application of MLR that describes the relationship between the response variable (length of the procedure) and the independent variables in this test is given by (5.2)

Length of the procedure = 1.0e-003*(0,2422 - 0,0009*Col_1 - 0,0001*Col_2 - 0,0005*Col_3 - 0,0005*Col_4 - 0,0001*Col_5 - 0,0002*Col_6 + 0,0002*Col_7- 0,0001*Col_8 + 0,0007*Col_9 - 0,0003*Col_10 -0,0001*Col_11 + 0,0001*Col_12 + 0,0005*Col_13 - 0,0002*Col_14 - 0,0002*Col_15 - 0,0003*Col_16 + 0,0001*Col_17- 0,0003*Col_18 + 0,0003*Col_19 + 0,0003*Col_20 -0,0001*Col_21 + 0,0001*Col_22 + 0,0009*Col_23) (5.2)

When applying the regression equation to the same training and test sets used in RNAs the MSE obtained (as described in equation 4.1) was equal to 0.0743 for the training set, while in test set error it was of 0.1287.

The Decision Tree technique was applied from the software WEKA (Waikato Environment for Knowledge Analysis), which is free and has an open code source, used for data mining. As in the MLR, when applying the decision tree technique it was used the same data sets and also the three types of tests were carried out. This technique also showed the best results in the 3rd test, where the average quadratic error was of 0.0881.

6 CONCLUSIONS

The Forum of Labor SJP has increased considerably the number of labor lawsuits. Given this fact, it is necessary to use mathematical optimization tools such as, from the area of Operational Research, which might in some way assist the legal department in its various procedures. In this work, these tools were used in order to enable the "negotiation" between the parties, by predicting the length of the labor lawsuits' proceedings.

The application presented here, related to processes of the Labor Court, shows, once again, the wide applicability of the techniques from the field of Operational Research. The application discussed here, aims to compare the techniques of ANNs, MLR and Decision Tree to find the best prediction. With data from 100 cases, which are the inputs to the techniques, we sought to obtain, automatically, a length forecast of the steps of the processes.

The ANNs were trained through the back-propagation algorithm, by the elaboration of a program using Visual Basic 6.0 software, varying the possibility of encoding the attributes, the number of neurons in the hidden layer, the set of initial weights and the number of iterations. The best response obtained showed an error of 0.07108 to an ANN with 23 neurons in the input layer, six neurons in the hidden layer, with 50 iterations (Table 5.2).

The MLR was performed using STATIGRAPHICS Plus 5.1 Software. In tests with this tool, the data sets used (training and testing) were the same of the ANNs, in order to obtain comparative parameters between the two mathematical tools. The error obtained was equal to 0.0743 for the training set and 0.1287 for the test set.

The Decision Tree technique was applied via WEKA software (Waikato Environment for Knowledge Analysis), which is free and has an open code source, used for data mining. We used the same sets of data in order to compare the applied techniques. With this tool the error obtained was 0.0881. With this tool the 3rd test also showed the best result, considering a full MSE equal to 0.0881.

Although the techniques have shown satisfactory results, the ANNs presented a superior performance when compared to other methods, as it can observed through the errors 0.07108 (ANNs), 0.1287 (MLR) and 0.0881 (Decision Tree).

Thus, the best way to predict the length of the processing of a new labor lawsuit, is to use ANN "E23N6I50", Table 5.2, where the weights were generated by the 3 third test (ordinal attributes without encoding, and with PCA) . This way, when it is desired to know the length of proceeding of a new labor lawsuit, one must determine the principal components of this case and then, using the topology and network weights E23N6I50, obtain the required number of months for this case.

It is worth mentioning that from time to time, in accordance with the suggestion of specialists of the area (labor judges), latest data (files) with known and reliable answers should be included in the database and methodology should be repeated, always glimpsing the lowest possible error for that prediction. With this, it is expected to obtain a more dynamic and accurate judiciary system, as well as greater satisfaction of its users.

REFERENCES

Adamowicz, E. C., 2000. Reconhecimento de Padrões na Análise Econômico–Financeira de Empresas. Curitiba. Dissertação de Mestrado, *PPGMNE, UFPR*.

Ambrósio, P. E., 2002 Redes Neurais Artificiais no Apoio ao Diagnóstico Diferencial de Lesões Intersticiais Pulmonares. Ribeirão Preto. *Dissertação de Mestrado, USP*.

Baesens, B., Setiono, R., Mues, C. & Vanthienen, J., 2003. Using Neural Network Rule Extraction and Decision Tables for Credit-Risk Evalution. *Management Science, 49*, 3, 312-329.

Baptistella, M., Cunico, L. H. B., Steiner, M. T. A., 2009. O Uso de Redes Neurais na Engenharia de Avaliações: Determinação dos Valores Venais de Imóveis Urbanos. *Revista de Ciências Exatas e Naturais*, 9, 2, 215-229.

Biondi Neto, L., Sieira, A. C. C. F., Danziger B. R., Silva, J. G. S., 2006. Neuro-CPT: Classificação de Solos usando-se Redes Neurais Artificiais. *Engevista, v. 8,* p. 37-48.

Bond, M. T., Seiler, V. L., Seiler, M. J., 2002. Residencial Real Estate Prices: a Room with a View. *The Journal of Real Estate Research*, v. 23, n. 1, p. 129-137.

Haykin, S., 2002. Redes Neurais: Princípios e Prática. Bookman, Porto Alegre, RS.

Lima, J. D., 2002. Análise Econômico–Financeira de Empresa Sob a Ótica da Estatística Multivariada. Curitiba. *Dissertação de Mestrado, PPGMNE, UFPR.*

Lu, H.; Setiono, R. & Liu, H., 1996. Effective Data Mining Using Neural Networks. *IEE Transactions on Knowledge an Data Engineering*, 8, 6, 957-961.

Nguyen, N., Cripps, A. 2001. Predicting Housing Value: A Comparison of Multiple Regression Analysis and Artificial Neural Networks. *The Journal of Real Estate Research*, v. 22, n. 3, p. 313-336.

SLC - Superior Labor Court. (http://www.tst.gov.br/) 16 february 2007.

Sousa, E. A., Teixeira, L. C. V., Mello, M. R. P. A., Torres, E. A. F. S., Moita Neto, J. M., 2003. Aplicação de Redes Neurais para Avaliação do Teor de Carne Mecanicamente Separada em Salsicha de Frango. *Ciência e Tecnologia de Alimentos*, 23, 3, Campinas.

Steiner, M. T. A. 1995. Uma Metodologia Para o Reconhecimento de Padrões Multivariados com Resposta Dicotômica. Florianópolis. Tese de Doutorado, Programa de Pós Graduação em *Engenharia de Produção*, UFSC.

Steiner, M. T. A., Nievola, J. C., Soma, N. Y., Shimizu, T., Steiner Neto, P. J., 2007. Extração de regras de classificação a partir de redes neurais para auxílio à tomada de decisão na concessão de crédito bancário. *Revista Pesquisa Operacional*, 27, 407-426.

Steiner, M. T. A.; Mota, J. F., 2007. Estudando um Caso de determinação do Preço de Venda de Imóveis Urbanos utilizando Redes Neurais Artificiais e Métodos Estatísticos Multivariados. *X Encontro de Modelagem Matemática, Nova Friburgo, RJ.*

Steiner, M. T. A., Bráulio, S. N., Alves, V., 2008. Métodos Estatísticos Multivariados aplicados à Engenharia de Avaliações, *Revista Gestão & Produção*, 15, 23-32.

Steiner, M. T. A., Carnieri, C., Stange, P., 2009. Construção de um Modelo Matemático para o Controle do Processo de Produção do Papel Industrial. *Pesquisa Operacional para o Desenvolvimento*, 1, 1, 33-49.

TRT - Tribunal Regional do Trabalho. (http://www.trt9.gov.br/> 07 october 2007.

Witten, I. H., Frank, E., 2005, Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann Publishers, 2nd edition**.**