

## TÉCNICAS DE RECONHECIMENTO DE PADRÕES APLICADAS À JUSTIÇA DO TRABALHO

**Alessandra Memari Pavanelli<sup>1</sup>; Genival Pavanelli<sup>2</sup>; Maria Teresinha Arns Steiner<sup>3</sup>;  
Deise Maria Bertholdi Costa<sup>4</sup>; Dr. Bráulio Gabriel Gusmão<sup>5</sup>**

<sup>1,2,3,4</sup>Universidade Federal do Paraná (UFPR); <sup>1,2,3,4</sup>Programa de Pós-Graduação em Métodos Numéricos em Engenharia (PPGMNE); <sup>3,4</sup>Programa de Pós-Graduação em Engenharia de Produção PPGEPE; <sup>4</sup>Departamento de Expressão Gráfica;

CP 19081 – Curitiba, PR; CEP 81531-990

<sup>3</sup>Pontifícia Universidade Católica do Paraná (PUCPR); <sup>3</sup>Programa de Pós-Graduação em Engenharia de Produção e Sistemas (PPGEPS);

<sup>5</sup>Juiz do Trabalho do Fórum Trabalhista de São José dos Pinhais, Paraná

<sup>1</sup>[alessandracmc@bol.com.br](mailto:alessandracmc@bol.com.br); <sup>2</sup>[pavanelllicmc@bol.com.br](mailto:pavanelllicmc@bol.com.br); <sup>3</sup>[tere@ufpr.br](mailto:tere@ufpr.br) ;  
<sup>4</sup>[deise@ufpr.br](mailto:deise@ufpr.br); <sup>4</sup>[brauliofusmao@trt9.gov.br](mailto:brauliofusmao@trt9.gov.br)

### Resumo

O presente trabalho tem como objetivos prever o tempo de duração de audiências trabalhistas e, também, o tempo de duração do trâmite de processos trabalhistas aos usuários do sistema de justiça. Desta forma, pretende-se otimizar tanto o agendamento dos horários dos juízes, através da geração de uma “agenda inteligente”, assim como, a “negociação” entre as partes envolvidas nos processos. A metodologia proposta consiste em aplicar e comparar duas técnicas da área da Pesquisa Operacional, Redes Neurais Artificiais (RNAs) e Regressão Linear Múltipla (RLM), com o intuito de obter o melhor desempenho possível para as previsões citadas. Para tanto, utilizou-se os dados de processos trabalhistas do Fórum Trabalhista de São José dos Pinhais, Paraná, para se fazer o treinamento de diversas RNAs com várias topologias e, também, da RLM. Em diversas simulações, ambas as técnicas foram utilizadas diretamente e, em outras, a Análise de Componentes Principais (ACP) e/ou a codificação dos atributos foi executada preliminarmente a utilização das mesmas, visando melhorar ainda mais os seus desempenhos. Assim, obtendo-se novos dados (processos) para os quais seja necessária qualquer uma das previsões, tenha-se condições de “diagnosticar” os seus tempos (duração da audiência ou duração do trâmite do processo), preliminarmente as suas execuções. As duas técnicas utilizadas mostraram-se eficientes, apresentando resultados consistentes com uma margem de erro aceitável.

**Palavras-chave:** Programação Matemática; Redes Neurais Artificiais; Regressão Linear Múltipla; Análise de Componentes Principais; Codificação dos Atributos.

### Abstract

This research aims at predicting the length of time of labor hearings and to predict the process length of time to the users of this law system in order to optimize the judges' time schedules in São José dos Pinhais Labor Court in Paraná State. To achieve the best result, the methodology used here is a comparison of two distinct techniques: Artificial Neural Network (ANN) and Multiple Linear Regression (MLR). Data collected from São José dos Pinhais Labor Court fed both ANNs and MLR tests. Many ANNs were trained using different topologies, or even using Principal Component Analysis with the objective of reaching improved results. By obtaining new necessary data for any forecast one will be able to ‘diagnose’ its timings (length of time hearings or process) before its execution. Both tools employed in this research have shown to be efficient, once they have presented consistent results with low error margin.

**Keywords:** Mathematical Programming, Artificial Neural Networks, Multiple Linear Regression; Principal Component Analysis; Attributes Coding.

### 1. Introdução

Este trabalho apresenta duas propostas de aplicação de técnicas da área da Pesquisa Operacional, junto à justiça do trabalho. A primeira aplicação consiste em estimar o tempo aproximado de cada audiência em um fórum trabalhista, mais especificamente, o Fórum Trabalhista de São José do Pinhais, PR, a fim de gerar uma “agenda inteligente” para as mesmas. A outra é fornecer uma previsão do tempo de duração de um processo trabalhista aos usuários do sistema da Justiça do Trabalho.

Tanto na primeira, quanto na segunda proposta, utilizou-se dois métodos: um da área de inteligência artificial, Redes Neurais Artificiais (RNAs) e outro, da área estatística, Regressão Linear Múltipla (RLM). As RNAs implementadas neste trabalho são do tipo *feed-forward*, as quais foram treinadas através do algoritmo *back-propagation*, utilizando a função de transferência sigmoideal. O objetivo de utilizar estes dois métodos já consagrados na literatura, é o de se fazer uma comparação entre os resultados finais obtidos e, com isso, verificar qual fornece o melhor desempenho (maior percentual de acertos) e, desta forma, ser usado em futuras previsões.

O presente trabalho está estruturado da seguinte forma: na seção 2 estão apresentados trabalhos correlatos que também fizeram uso das técnicas da Pesquisa Operacional aqui aplicadas. Na seção 3 está a descrição do problema, levantamento e tratamento dos dados. Na seção 4 está a metodologia do trabalho, onde são apresentados os conceitos que envolvem as técnicas de RNAs, Análise de Componentes Principais (ACP) e RLM. A seção 5 descreve a implementação computacional das técnicas utilizadas e a análise dos resultados. Finalizando, na seção 6 são apresentadas as conclusões obtidas através da análise dos resultados da seção anterior.

### 2. Trabalhos Correlatos

Existem na literatura, numerosos trabalhos relacionados à previsão de dados, nos quais várias técnicas da área de Pesquisa Operacional e, mais especificamente, de Reconhecimento de Padrões, têm sido aplicadas. Vale salientar que não foram encontrados trabalhos de previsão relacionados a problemas da Justiça do Trabalho, como o aqui apresentado. Dentre os trabalhos pesquisados na literatura, pode-se destacar os relacionados a seguir.

Em BAPTISTELLA, CUNICO e STEINER (2009), os autores procuram por técnicas alternativas com o objetivo de determinar valores venais para imóveis de Guarapuava, PR. É proposta a utilização das RNAs e, para isso, foram coletados 256 registros históricos (padrões) de imóveis urbanos da cidade. Cada um dos registros ficou composto por 13 informações (atributos): bairro, setor, pavimentação, esgoto, iluminação pública, área do terreno, pedologia, topografia, situação, área edificada, tipo, estrutura e conservação. Várias simulações foram desenvolvidas, sendo que os piores resultados apresentaram acurácia de 78% e os melhores, de 95%.

Ainda na avaliação imobiliária, tem-se ainda o trabalho de NGUYEN e CRIPPS (2001), que compara o desempenho das RNAs com a Análise de Regressão Múltipla para a venda de casas de família. Múltiplas comparações foram feitas entre os dois modelos nas quais foram variados: o tamanho da amostra de dados, a especificação funcional e a predição temporal. No trabalho de BOND; SEILER e SEILER (2002), os autores examinam o efeito que a vista de um lago (Lago Erie, E.U.A.) tem sobre o valor de uma casa. No estudo foram levados em consideração os preços baseados na transação das casas (preço de mercado). Os resultados indicam que, além da variável vista, que se apresenta significativamente mais importante do que as demais, também a área construída e o tamanho do lote são importantes.

## PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

BAESENS *et al.* (2003) abordam três métodos para a extração de regras de uma rede neural, comparativamente: *Neurorule*; *Trepan* e *Nefclass*. Para comparar os desempenhos dos métodos abordados, foram utilizadas três bases de dados reais de crédito: *German Credit* (obtida do repositório *UCI*), *Bene 1* e *Bene2* (obtidas das duas maiores instituições financeiras da *Benelux*). Os algoritmos mencionados são ainda comparados com os algoritmos *C4.5-árvore*, *C4.5-regras* e Regressão Logística. Os autores ainda mostram como as regras extraídas podem ser visualizadas como uma tabela de decisão na forma de um gráfico compacto e intuitivo, permitindo uma melhor leitura e interpretação dos resultados ao gerente de crédito.

Em STEINER *et al.*, 2007, os autores apresentam uma metodologia, composta por técnicas de Análise Multivariada, para a construção de um modelo estatístico de Regressão Linear Múltipla para avaliação de imóveis também. É aplicada, inicialmente, a Análise de Agrupamento aos dados de cada classe de imóvel urbano (apartamentos, residências e terrenos) para obtenção de grupos homogêneos dentro de cada classe e, em correspondência, são determinados discriminantes para alocar futuros itens nesses grupos, pelo Método do Escore Discriminante Quadrático. Em seguida, é aplicada a técnica da ACP para resolver o problema da multicolinearidade que pode existir entre as variáveis do modelo. Com os escores das componentes principais é ajustado um modelo de Regressão Linear Múltipla para cada grupo de imóveis homogêneos dentro de cada classe. A metodologia foi aplicada a um conjunto de 119 imóveis (44 apartamentos, 51 casas e 24 terrenos), da cidade de Campo Mourão, PR. O modelo para cada grupo homogêneo dentro de cada classe de imóveis avaliados apresentou um ajuste adequado aos dados e uma capacidade preditiva bastante satisfatória.

Já ADAMOWICZ (2000) utiliza técnicas de Reconhecimento de Padrões, RNAs e Análise Discriminante Linear de Fisher, com o objetivo de classificar empresas como solventes ou insolventes. Os dados foram fornecidos pelo Banco Regional de Desenvolvimento do Extremo Sul (BRDE), Regional de Curitiba, PR. As duas técnicas se apresentaram eficientes na discriminação das empresas, sendo que o desempenho das RNAs se apresentou um pouco melhor que a Análise Discriminante Linear de Fisher.

AMBROSIO (2002) apresenta um trabalho que tem como objetivo desenvolver um sistema computacional para auxiliar médicos radiologistas na confirmação de diagnósticos das chamadas lesões intersticiais pulmonares. Os dados foram obtidos junto ao Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (HCFMRP) por meio de protocolos gerados por especialistas. O sistema foi desenvolvido com o uso de RNAs de múltiplas camadas como um classificador de padrões. O algoritmo de treinamento foi o *back-propagation* com a função de ativação sigmoideal. Vários testes foram realizados para diferentes configurações de redes. Ficou claro que o uso desta ferramenta é viável, já que uma vez a rede treinada e seus pesos fixados, não é preciso mais acessar o banco de dados. Isso torna o sistema mais rápido e computacionalmente mais leve. A pesquisa conclui que as RNAs cumprem bem seu papel como classificadores de padrões.

SOUSA *et al.* (2003) utilizaram técnicas de RNAs com três camadas de neurônios com o algoritmo *back-propagation*. O objetivo foi fazer a previsão do teor de carne mecanicamente separada (CMS) em produtos cárneos, a partir do teor de minerais contidos nas salsichas formuladas com diferentes teores de carne de frango. A técnica utilizada mostrou-se bem eficiente durante a fase de treinamento e testes, no entanto, a aplicação das RNAs às amostras comerciais foi inadequada, devido à diferença de ingredientes das salsichas usadas no treinamento e os ingredientes das amostras comerciais.

Em STEINER, CARNIERI e STANGE (2009), é proposta a utilização de um modelo de Programação Linear para o Reconhecimento de Padrões de bobinas de papel de boa ou baixa

qualidade. Foram coletados dados de 145 bobinas de papel (padrões), 40 de boa qualidade e 105 de baixa qualidade. De cada bobina foram considerados 18 atributos: testes de tração e rasgo da celulose, da pasta mecânica e da pasta termo-mecânica; quantidades destas três pastas; consistência e vazão da celulose e dados de sete rolos de prensagem da máquina de papel. A partir do modelo de PL, foi construído um 2º. modelo matemático que faz uso do 1º., de forma a garantir a obtenção de bobinas de boa qualidade a um mínimo custo.

BIONDI NETO *et al.* (2006) mostram em seu trabalho que a determinação do tipo de solo, até então, podia ser obtida através de ábacos; o objetivo da pesquisa foi aplicar um método computacional para classificar o solo. A técnica utilizada foi novamente a de RNAs com método de *levenberg-marquardt*, que apresenta como resultado a classificação do solo para cada incremento de profundidade. Todos os dados foram obtidos de situações reais. O tempo de convergência foi rápido, o que facilitou a realização de vários testes.

LU; SETIONO e LIU (1995; 1996) relatam em seus artigos o algoritmo chamado *Neurorule* que faz a extração de regras a partir de uma rede neural treinada, obtendo regras do tipo SE-ENTÃO ("*IF-THEN*"). O desempenho desta abordagem é verificado, em ambos os artigos, em um problema de crédito bancário, sendo que para facilitar a referida extração de regras, os valores dos atributos numéricos foram discretizados, dividindo-os em sub-intervalos. Após a discretização, o esquema de codificação "termômetro" foi empregado para obter representações binárias dos intervalos anteriormente definidos obtendo-se, assim, as entradas para a rede neural. Os resultados obtidos nos artigos indicam que, usando a abordagem proposta, regras de alta qualidade podem ser descobertas a partir de um conjunto de dados.

E assim, vários outros trabalhos das mais diversas áreas de pesquisa, fazendo uso das mais diversas técnicas de Reconhecimento de Padrões, em especial das RNAs, poderiam aqui ser citados.

### 3. Descrição do Problema, Levantamento e Codificação dos Atributos

Atualmente, muitos países possuem legislação trabalhista, mas nem sempre foi assim. No Brasil, a Justiça do Trabalho e a legislação trabalhista surgiram somente depois do século XIX, após várias lutas e reivindicações provenientes das classes operárias. Somente a partir da Revolução de 1930 é que foi criado o Ministério do Trabalho, e, a Justiça do Trabalho foi prevista pela Constituição de 1934. Atualmente a Justiça do Trabalho está estruturada em três graus de jurisdição:

- Primeiro Grau: Varas do Trabalho;
- Segundo Grau: Tribunais Regionais do Trabalho;
- Terceiro Grau: Tribunal Superior do Trabalho.

De acordo com o Tribunal Superior do Trabalho (TST), existem em todo o Brasil 24 Tribunais Regionais do Trabalho (TRT), sendo que a partir de 2003, cerca de 270 novas Varas do Trabalho foram criadas, com o objetivo de acelerar os trâmites legais dos processos trabalhistas (TST, 2007). Só no estado do Paraná, no TRT da 9ª Região, existem 28 desembargadores distribuídos em todo o estado (TRT, 2007). Das 77 Varas do Trabalho do estado do Paraná, a de São José dos Pinhais (SJP) ocupa o 2º lugar em número de ações trabalhistas. No ano de 2006, o Fórum Trabalhista de SJP passou a contar com a 2ª Vara do Trabalho. Devido ao aumento do número de processos trabalhistas, como consequência da industrialização maciça no município, faz-se necessário a agilidade nos serviços da justiça. Com isso, o uso de ferramentas matemáticas para o "agendamento inteligente" das audiências, bem como a previsão do tempo de duração do trâmite dos processos trabalhistas, é de fundamental importância para essa agilização.

## PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

Os dados dos processos (padrões), assim como os atributos de cada padrão, utilizados para o desenvolvimento do presente trabalho, foram obtidos junto a 1ª Vara do Trabalho do Fórum de SJP, PR. Com o objetivo de determinar quais atributos seriam relevantes na determinação da duração de uma audiência trabalhista e, também, na determinação do tempo de duração do trâmite de um processo trabalhista, foram realizadas várias reuniões com o juiz titular deste Fórum. Como resultado destas reuniões chegou-se ao conjunto de 12 atributos listado a seguir.

- a. Rito: que pode ser do tipo trabalhista (RT) ou um procedimento sumaríssimo (PS);
- b. Tempo de Serviço: é a diferença entre a data de dispensa e data de admissão, em meses;
- c. Salário do Reclamante: último salário recebido;
- d. Profissão: função exercida pelo reclamante. Este atributo foi dividido em duas partes: setor que, por sua vez, se subdivide em comércio, indústria e serviço; e cargo, que se subdivide em direção e execução;
- e. Objeto do Processo: corresponde às solicitações feitas pelo reclamante. Podem ser: falta de registro em carteira profissional, diferenças salariais, verbas rescisórias, multa do Art. 477, multa do Art. 467, horas extras e reflexos, fundo de garantia por tempo de serviço, indenização por danos morais, seguro desemprego, vale transporte, adicional de insalubridade, adicional noturno e plano de saúde;
- f. Juiz: a 1ª Vara do Trabalho conta com mais dois juízes substitutos além do juiz titular;
- g. Depoimentos: quantidade de depoimentos em cada audiência;
- h. Acordo: quando ocorre o acordo entre as partes;
- i. Perícia: se há necessidade ou não da realização de alguma espécie de perícia como, por exemplo, a perícia médica ou a de periculosidade;
- j. Recurso Ordinário: quando uma das partes (reclamante ou reclamado) não concorda com a sentença emitida pelo juiz e solicita Recurso Ordinário ao TRT;
- k. Recurso de Revista: quando uma das partes (reclamante ou reclamado) não concorda com o acórdão emitido pelo TRT e solicita Recurso de Revista ao TST;
- l. Número de Audiências: trata do número de audiências necessárias para que o juiz emita a sentença;

Foram coletados dados de 100 processos para as referidas previsões. Os 12 atributos listados anteriormente constituíram a matriz destinada ao treinamento e ao teste das RNAs, bem como para aplicação da técnica de RLM.

A maioria dos dados foi tratada de forma a corresponder a uma ou mais coordenadas binárias (LU; SETIONO e LIU, 1996), (BAESENS *et al.*, 2003) do vetor de entradas para as técnicas utilizadas, conforme comentado na seção 3.1, a seguir.

### 3.1 Codificação dos Atributos

A fim de se tentar melhorar o desempenho das técnicas utilizadas (RNAs e RLM), cada um dos 12 atributos citados foi "tratado" de maneira a corresponder a uma ou mais coordenadas binárias (LU; SETIONO e LIU, 1996), (BAESENS *et al.*, 2003), dependendo se o mesmo era nominal ou ordinal. Foi utilizada a "codificação termômetro" para os atributos ordinais e a "codificação *dummy*" (artificial) para os atributos nominais (BAESENS *et al.*, 2003), (STEINER *et al.*, 2007).

A tabela 3.1 ilustra a "codificação termômetro" para o atributo ordinal "Salário do Reclamante", por exemplo. Este atributo é, primeiramente, discretizado nos valores de 1 a 5; se, por exemplo, a "Entrada 1 = 1", isto significa que a variável original "Salário do Reclamante" > 1.340. A tabela 3.2 ilustra a "codificação *dummy*" para a variável nominal "Acordo", por exemplo.

**Tabela 3.1.** Um exemplo do procedimento "codificação termômetro" para variáveis ordinais.

Entrada original "Salário do Reclamante" SR(reais)	Entrada Categorica	Entrada 1	Entrada 2	Entrada 3	Entrada 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

**Tabela 3.2.** Um exemplo do procedimento "codificação *dummy*" para variáveis nominais.

Entrada original "Acordo"	Entrada 1
Acordo = sim	0
Acordo = não	1

Para a previsão do tempo de duração de audiências trabalhistas ("agenda inteligente") não foram utilizados os atributos: "Perícia", "Recurso Ordinário", "Recurso de Revista" e "Número de Audiências", uma vez que os mesmos não possuem relação com esta previsão. Sendo assim, o vetor de entrada dos dados para esta previsão possui 8 atributos, que codificados fornecem 30 coordenadas relativas a cada um dos processos trabalhistas, conforme detalhamento na tabela 1, em anexo. Deste modo, a matriz de entradas para esta previsão possui 100 linhas (processos; padrões) e 30 colunas (atributos codificados). Já para a previsão do tempo de duração do trâmite de um processo trabalhista não foi utilizado o atributo "Juiz", uma vez que durante o processo vários juizes podem atuar. Também não foram considerados "Depoimentos" e "Objeto do processo – plano de saúde". Assim, para este previsão têm-se 10 atributos, que codificados fornecem 32 atributos; logo, a matriz de entradas para esta previsão, possui 100 linhas e 32 colunas (ver tabela 1 em anexo).

#### 4. Metodologia do Trabalho

A metodologia aplicada neste trabalho buscou, por meio da utilização das RNAs e da RLM, comparativamente, reconhecer padrões nos processos trabalhistas analisados, visando prever o tempo de audiências trabalhistas e, também, o tempo do trâmite de processos trabalhistas aos usuários desse sistema de justiça, conforme já comentado.

##### 4.1 RNAs

A RNA implementada neste trabalho, classificada como Rede de Múltiplas Camadas ou Rede *feed-forward*, foi treinada através do algoritmo *back-propagation*, utilizando a função de transferência sigmoideal, que gera saídas entre "0" e "1" para entradas entre  $-\infty$  e  $+\infty$ , em todos os neurônios. O desempenho da rede foi verificado através do MSE (erro médio quadrático), dado pela equação (4.1).

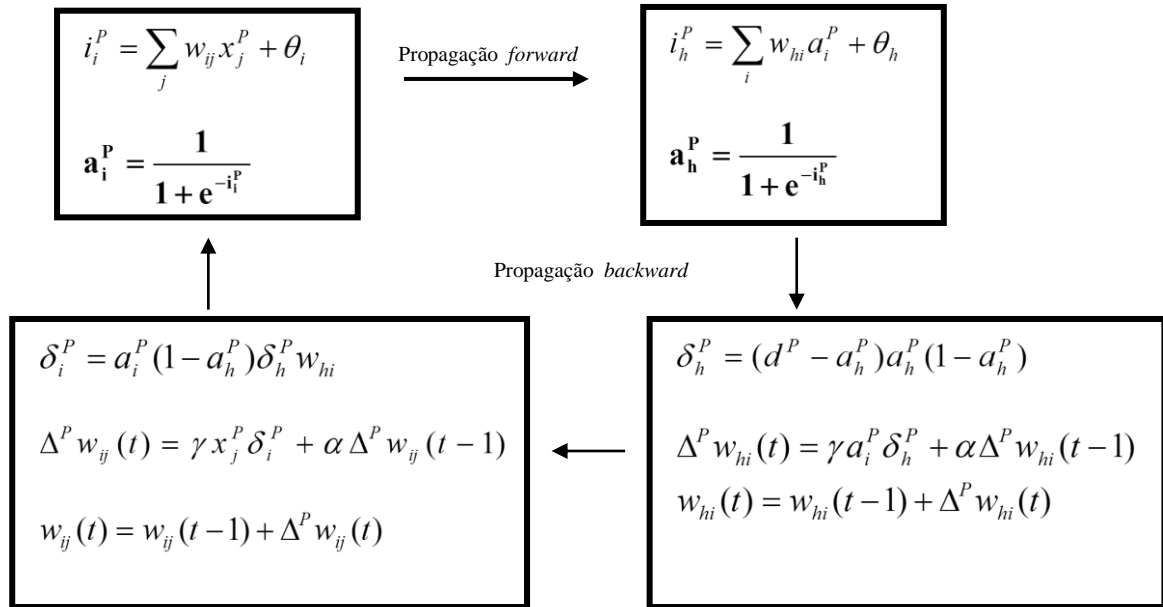
$$MSE = \frac{\sum_{i=1}^n (d_i^p - a_i^p)^2}{2n} \tag{4.1}$$

onde  $n$  = número de padrões,  $d_i^p$  é a saída desejada (valor real) para o padrão  $p$  e  $a_i^p$  é a saída obtida, pela rede, para o padrão  $p$ .



Na figura 4.1, a seguir, é apresentado um esquema do algoritmo implementado quando  $h=1$ , ou seja, quando a rede apresenta um único neurônio na camada de saída.

**Figura 4.1.** Ilustração do algoritmo *back-propagation*



Nesta Figura 4.1, tem-se:  $p$  = padrão;  $t$  = iteração;  $\gamma$  = taxa de aprendizagem;  $j$  = camada de entrada;  $i$  = camada oculta;  $h$  = camada de saída;  $\alpha$  = taxa de momento;  $w_{ij}$  e  $w_{hi}$  = pesos entre as camadas de entrada-oculta e oculta-saída, respectivamente;  $i^p$  = *input* para o padrão  $p$ ;  $x_j^p$  = entrada para o padrão  $p$  na camada  $j$  (de entrada);  $a^p$  = *output* para o padrão  $p$ ;  $d^p$  = valor desejado para o padrão  $p$ .

#### 4.2 RLM

O segundo método utilizado neste trabalho, RLM, tem como principal objetivo descrever a relação entre uma variável resposta e uma ou mais variáveis explicativas. Os tipos de regressão mais utilizados são: Logística e Linear, muito utilizadas em diversas áreas do conhecimento.

Segundo Lima, 2002, em 1845 surgiu a técnica Regressão Logística com o objetivo de resolver problemas de crescimento demográfico. Essa técnica passou também a ser empregada na área da Biologia na década de 30. Já sua aplicação em problemas econômicos e sociais aparece somente nos anos 60. Recentemente, essa metodologia tornou-se referência obrigatória em muitos manuais de econometria. A Regressão Logística é uma técnica estatística muito usada em análise de dados com respostas pertencentes ao intervalo  $[0, 1]$ , com o objetivo de classificar padrões em classes.

Já a Regressão Linear é muito utilizada em diversas áreas da pesquisa, sendo um tipo de técnica que pode produzir valores de resposta estimada fora do intervalo  $[0, 1]$ . É considerado um modelo clássico de regressão. É uma técnica utilizada para estudar a relação entre uma variável dependente e várias variáveis independentes. O objetivo pode ser explicativo, ou seja, demonstrar uma relação matemática que pode indicar, mas não provar, uma relação causa e efeito, ou preditivo, quer dizer, obter uma relação que permita, perante futuras observações das variáveis  $x_p$ , prever o correspondente valor de  $y$ .

## PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

Suponha que está se querendo construir um modelo que relacione a variável resposta  $y$  com  $p$  fatores  $x_1, x_2, \dots, x_p$ . Esse modelo sempre incluirá uma parcela de erro. Tem-se então:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon$$

para  $i = 1, 2, \dots, n$  onde  $n$  é o número de observações;  $p$  o número de variáveis.

Usando notação matricial:  $Y = X\beta + \varepsilon$ , onde  $Y$  é a variável resposta;  $X$  matriz do modelo;  $\beta$  é o vetor de parâmetros a serem estimados;  $\varepsilon$  vetor de erros aleatórios.

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

### 4.3 ACP

Buscando melhorar ainda mais os resultados obtidos, em alguns dos testes foi aplicada a ACP que é capaz de identificar padrões nos dados, visando expressar os mesmos de modo a salientar as similaridades e também diferenças existentes. Está relacionada com a explicação da estrutura da covariância por meio de poucas combinações lineares das variáveis originais.

Além disso, através da ACP reduz-se a dimensão original da matriz de dados através de combinações lineares de um conjunto de variáveis que retenham o máximo possível de informação contida nas variáveis originais e, ainda, facilita a interpretação das análises realizadas, julgando a importância das próprias variáveis originais escolhidas.

## 5. Implementação Computacional e Obtenção dos Resultados

Conforme descrito na seção 4 deste trabalho, os métodos propostos neste trabalho, RNAs e RLM, foram aplicados após a coleta e o tratamento (codificação dos atributos/aplicação da ACP) dos processos analisados, que foram ajuizados junto ao Fórum Trabalhista de SJP. Todos os dados obtidos de cada um dos processos trabalhistas serviram para compor as matrizes de entrada. O treinamento das RNAs implementado neste trabalho é do tipo supervisionado, ou seja, para cada vetor de entrada de dados já se conhece a saída (HAYKIN, 2001). Dessa forma, a fim de executar o treinamento e testes das RNAs foi implementado um programa utilizando o *Software Visual Basic 6.0*.

Para a realização do treinamento da RNAs foi utilizado, conforme já comentado, o algoritmo supervisionado *back-propagation*, com a função de ativação sigmoideal, com saídas no intervalo  $[-1, 1]$ . Devido a estas condições da função de ativação, foi necessário enquadrar as saídas neste intervalo. Assim, os tempos de duração de cada audiência, que variam de 1 a 93 minutos, foram divididos por 93, enquanto que os tempos de duração de trâmite dos processos, que variam de 1 a 94 meses, foram divididos por 94. Vale salientar que ambos os tempos possuem distribuição uniforme.

Para a avaliação da RNA, foi utilizado o procedimento *hould out*, ou seja, do conjunto dos processos cadastrados, 75% foram utilizados para o treinamento da rede e os 25% restantes



## PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

foram utilizados no teste. Na aplicação das RNAs foram utilizados, em todos os testes, quatro conjuntos de pesos iniciais.

Em todos os testes realizados com as RNAs, primeiro variou-se o número de neurônios da camada oculta de “1” até “15”, mantendo-se fixo o número de 50 iterações para cada uma das topologias, a fim de verificar qual o menor erro no grupo de teste. A arquitetura que forneceu o menor erro, voltou a ser treinada, agora variando-se o número de iterações, até o momento em que o erro no grupo de testes atingisse o mínimo. Desta forma, estar-se-á evitando *overfitting* da rede, ou seja, a RNA passaria a apresentar melhores resultados para o grupo de treinamento, porém perderia a capacidade de generalização.

Foi escolhida uma nomenclatura para cada topologia, a fim de representar, em seqüência, as seguintes características da RNA: quantidade de entradas, número de neurônios na camada escondida e número de iterações. Por exemplo, a rede “E32N3I40” é uma rede com “32” entradas, “3” neurônios na camada escondida e que foi treinada com “40” iterações.

### 5.1. Resultados da Previsão do Tempo de Duração de Audiências Trabalhistas

Para a previsão do tempo de duração de audiências trabalhistas, dentre os testes realizados através das RNAs, o que apresentou melhor resultado, ou seja, aquele que apresentou menor erro, foi o 2º teste, onde os atributos foram codificados (seção 3) e, em seguida, submetidos à ACP. A tabela 5.1, a seguir, mostra a variação do número de neurônios da RNA para este teste.

**Tabela 5.1.** Resultados das Simulações variando do Número de Neurônios na Camada Escondida.

SIMULAÇÃO	TOPOLOGIA	MSE Tr	MSE Tes	SIMULAÇÃO	TOPOLOGIA	MSE Tr	MSE Tes
1	E30N1I50	0,00368	0,01845	9	E30N9I50	0,00066	0,01376
2	E30N2I50	0,00244	0,02811	<b>10</b>	<b>E30N10I50</b>	<b>0,00046</b>	<b>0,00562</b>
3	E30N3I50	0,00095	0,03154	11	E30N11I50	0,00050	0,02122
4	E30N4I50	0,00082	0,07066	12	E30N12I50	0,00062	0,02392
5	E30N5I50	0,00072	0,03289	13	E30N13I50	0,00051	0,02115
6	E30N6I50	0,00052	0,01832	14	E30N14I50	0,00056	0,03212
7	E30N7I50	0,00051	0,07542	15	E30N15I50	0,00041	0,02349
8	E30N8I50	0,00065	0,05511				

Nesta tabela 5.1 é mostrado para cada uma das 15 simulações executadas, o MSE para o conjunto de treinamento (MSE Tr) e o MSE para o conjunto de teste (MSE Tes). Verifica-se a partir desta tabela, que a melhor topologia, ou seja, que apresentou o menor erro para o MSE Tes, caracterizando uma melhor capacidade de generalização, foi a rede E30N10I50, ou seja, com 30 neurônios na camada de entrada, 10 neurônios na camada escondida e treinada com 50 iterações. O MSE para esta rede é de 0,00046 para o conjunto de treinamento e de 0,00562 para o conjunto de teste.

## PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

Em seguida, esta RNA (E30N10I50), voltou a ser treinada, só que desta vez, fazendo-se variar o número de iterações, buscando diminuir ainda mais o erro no conjunto de teste. Os resultados relevantes são apresentados na tabela 5.2 a seguir.

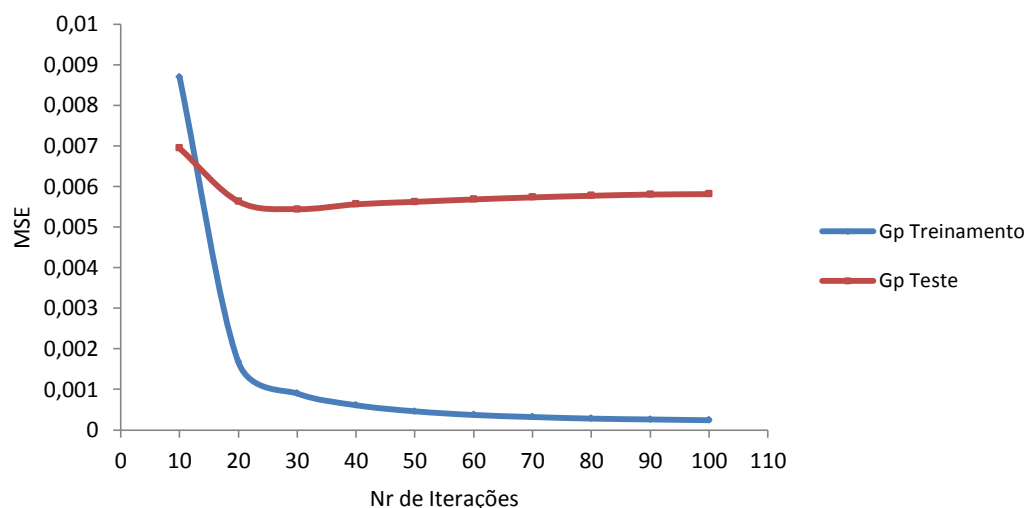
**Tabela 5.2.** Resultados das Simulações variando o Número de Iterações.

SIMULAÇÃO	TOPOLOGIA	MSE Tr	MSE Tes	SIMULAÇÃO	TOPOLOGIA	MSE Tr	MSE Tes
1	E30N10I10	0,00869	0,00694	9	E30N10I90	0,00026	0,00580
2	E30N10I20	0,00167	0,00563	10	E30N10I100	0,00024	0,00581
<b>3</b>	<b>E30N10I30</b>	<b>0,00090</b>	<b>0,00544</b>	11	E30N10I200	0,00014	0,00581
4	E30N10I40	0,00061	0,00556	12	E30N10I300	0,00009	0,00635
5	E30N10I50	0,00046	0,00562	13	E30N10I400	0,00006	0,00714
6	E30N10I60	0,00037	0,00568	14	E30N10I500	0,00004	0,00762
7	E30N10I70	0,00032	0,00573	15	E30N10I1000	0,00001	0,00857
8	E30N10I80	0,00028	0,00577				

O melhor resultado, conforme se pode verificar nesta Tabela 5.2, é obtido na 3ª. simulação, cuja topologia é E30N10I30, onde se tem 30 iterações e MSE igual a 0,00544 no grupo de teste.

O gráfico 5.1, a seguir, mostra como varia o erro tanto no grupo de treinamento como no grupo de teste. Observa-se que no grupo de treinamento o MSE diminui monotonamente a medida que se aumenta o número de iterações, enquanto que no grupo de teste o erro diminui até atingir o menor valor quando a rede foi submetida a 30 iterações. A partir dessa quantidade de iterações, se continuarmos treinando a rede, ou seja, ao aumentarmos o número de iterações o erro nesse grupo começa a aumentar novamente indicando que essa RNA perde a sua capacidade de generalização.

**Gráfico 5.1** Treinamento da RNA do Tempo de Duração de Audiência Trabalhista



Ainda, para a previsão deste tempo de audiência, agora através da técnica de RLM, utilizou-se os mesmos conjuntos de dados (treinamento e teste) das RNAs. As tabelas de dados

## PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

foram construídas no *Software Excel* e, para a obtenção dos resultados foi utilizado o *software STATIGRAPHICS Plus 5.1*.

Assim como nas RNAs, esta ferramenta foi aplicada nos três tipos de testes com o objetivo de comparar os resultados obtidos. O melhor resultado obtido com esta técnica, também foi com o 2º teste, onde os atributos foram codificados e, em seguida, submetidos à ACP. A equação de RLM que descreve a relação entre a variável resposta (tempo de audiência) e as variáveis independentes nesse teste é descrita a seguir, em (5.1).

$$\begin{aligned} \text{Tempo de audiência} = & 0,220922 - 0,0294401*\text{Col}_1 + 0,151429*\text{Col}_2 - 0,0265103*\text{Col}_3 + \\ & 0,134604*\text{Col}_4 - 0,0273278*\text{Col}_5 - 0,170564*\text{Col}_6 - 0,0816318*\text{Col}_7 + 0,11438*\text{Col}_8 - \\ & 0,065881*\text{Col}_9 - 0,0339234*\text{Col}_{10} - 0,00485417*\text{Col}_{11} + 0,117835*\text{Col}_{12} + \\ & 0,189141*\text{Col}_{13} - 0,0111986*\text{Col}_{14} - 0,065423*\text{Col}_{15} - 0,164584*\text{Col}_{16} - \\ & 0,186351*\text{Col}_{17} + 0,112364*\text{Col}_{18} + 0,0637656*\text{Col}_{19} - 0,312607*\text{Col}_{20} - \\ & 0,163999*\text{Col}_{21} - 0,132798*\text{Col}_{22} + 0,173492*\text{Col}_{23} + 0,116546*\text{Col}_{24} + \\ & 0,133817*\text{Col}_{25} - 0,0963288*\text{Col}_{26} + 0,817548*\text{Col}_{27} + 0,127331*\text{Col}_{28} - \\ & 0,0334636*\text{Col}_{29} - 7,25014\text{E}14*\text{Col}_{30} \end{aligned} \quad (5.1)$$

Ao se aplicar a equação (5.1), aos 100 processos trabalhistas, obteve-se um erro quadrático médio igual a 0,011215.

### 5.2. Resultados da Previsão do Tempo de Duração do Trâmite de Processos Trabalhistas

Na previsão do tempo de duração de trâmite de processos trabalhistas, o 3º teste foi o que apresentou melhor resultado, onde os atributos ordinais não foram codificados e, em seguida, os dados foram submetidos à ACP. A tabela 5.3, a seguir, mostra a variação do número de neurônios nas RNAs para este teste.

**Tabela 5.3.** Resultados das Simulações variando o Número de Neurônios na Camada Escondida.

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

De acordo com a tabela 5.3, verifica-se que a melhor topologia de rede é a E23N6I50, ou seja, 23 neurônios na camada de entrada e seis na camada escondida, treinada com 50 iterações. A partir desta análise, esta rede foi treinada, agora variando-se o número de iterações, a fim de se obter o menor erro possível. A tabela 5.4, a seguir, mostra os resultados relevantes.

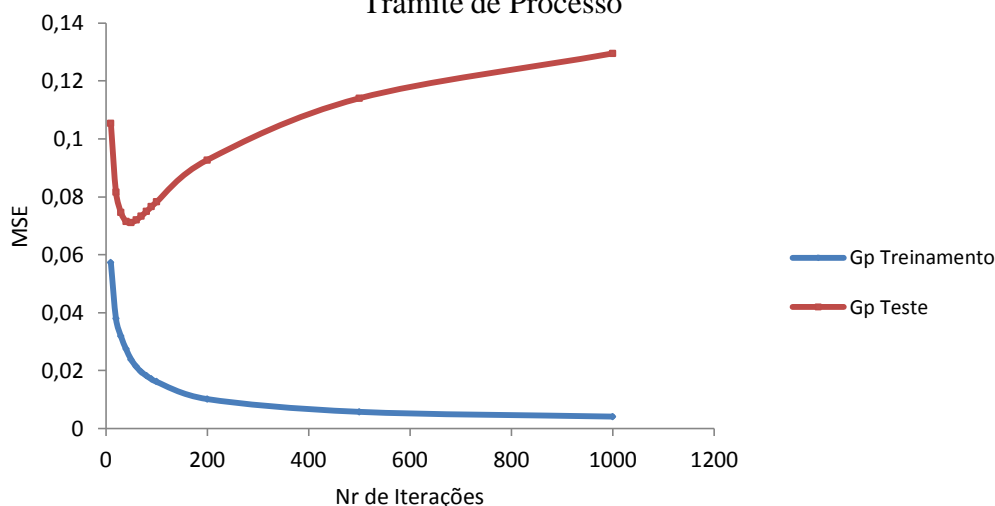
**Tabela 5.4.** Resultados das Simulações variando o Número de Iterações.

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

A partir dos resultados da tabela 5.4, é possível observar que a simulação com 50 iterações, já fornece o melhor resultado (menor erro no grupo de testes).

Como se esperava o erro no grupo de treinamento decresce monotonamente à medida que aumenta o número de iterações, já no grupo de testes o erro diminui atingindo um mínimo de 0,07108 quando a rede está sendo treinada com 50 iterações. Aumentando esse número nota-se claramente que o erro nesse grupo começa aumentar caracterizando a perda da capacidade de generalização da RNA a partir desse momento. Tais informações podem ser observadas no gráfico 5.2 a seguir.

**Gráfico 5.2** Treinamento da RNA do Tempo de Duração de Trâmite de Processo



Para esta mesma previsão (tempo do trâmite de processos), através da técnica de RLM, utilizou-se os mesmos conjuntos de dados (treinamento e teste) das RNAs, bem como os três tipos de testes com o objetivo de comparar os resultados obtidos.

O melhor resultado obtido com esta técnica, também foi com o 3º teste. O vetor de parâmetros estimados obtidos na aplicação da RLM que descreve a relação entre a variável resposta (tempo de trâmite) e as variáveis independentes nesse teste é dado por (5.2)

## PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

$$\begin{aligned} \text{Tempo de processo} = & 1.0e-003*(0,2422 - 0,0009*\text{Col}_1 - 0,0001*\text{Col}_2 - 0,0005*\text{Col}_3 - \\ & 0,0005*\text{Col}_4 - 0,0001*\text{Col}_5 - 0,0002*\text{Col}_6 + 0,0002*\text{Col}_7 - 0,0001*\text{Col}_8 + 0,0007*\text{Col}_9 - \\ & 0,0003*\text{Col}_{10} - 0,0001*\text{Col}_{11} + 0,0001*\text{Col}_{12} + 0,0005*\text{Col}_{13} - 0,0002*\text{Col}_{14} - \\ & 0,0002*\text{Col}_{15} - 0,0003*\text{Col}_{16} + 0,0001*\text{Col}_{17} - 0,0003*\text{Col}_{18} + 0,0003*\text{Col}_{19} + \\ & 0,0003*\text{Col}_{20} - 0,0001*\text{Col}_{21} + 0,0001*\text{Col}_{22} + 0,0009*\text{Col}_{23}) \end{aligned} \quad (5.2)$$

Ao se aplicar a equação de regressão linear aos 100 processos trabalhistas, obteve-se um erro quadrático médio igual a 0,0879.

### 6. Conclusões

O Fórum Trabalhista de SJP vem aumentando de maneira considerável o número de processos trabalhistas. Diante desse fato, faz-se necessária a utilização de ferramentas matemáticas de otimização como, por exemplo, da área de Pesquisa Operacional, que possam de alguma forma auxiliar a área jurídica em seus diversos procedimentos. Neste trabalho, tais ferramentas foram utilizadas com dois objetivos: criar uma “agenda inteligente”, através da previsão do tempo de audiências e, também, possibilitar a “negociação” entre as partes, através da previsão da duração do trâmite de processos trabalhistas.

As aplicações aqui apresentadas, relacionadas a processos da Justiça do Trabalho, mostram, mais uma vez, a grande aplicabilidade das técnicas da área de Pesquisa Operacional. As duas aplicações, aqui abordadas, visam comparar as técnicas de RNAs e RLM para encontrar a melhor previsão para ambos os casos estudados. Com dados obtidos de 100 processos, que são as entradas para ambas as técnicas e para ambas as aplicações, buscou-se obter, de forma automática, uma previsão de duração das audiências e dos trâmites dos processos.

As RNAs foram treinadas através do algoritmo *back-propagation*, por meio da elaboração de um programa utilizando o *Software Visual Basic 6.0*, fazendo-se variar a possibilidade de se codificar os atributos, o número de neurônios na camada escondida, o conjunto de pesos iniciais e o número de iterações. Para a melhor resposta (saída da RNA) obtida para a 1ª aplicação (tempo das audiências), o erro (MSE Tes) foi de 0,00544, para uma RNA com 30 neurônios na camada de entrada, 10 neurônios na camada oculta, com 30 iterações (tabela 5.2). Já para a 2ª aplicação (duração do trâmite do processo trabalhista), a melhor resposta obtida apresentou um erro de 0,07108, para uma RNA com 23 neurônios na camada de entrada, 6 neurônios na camada oculta, com 50 iterações (tabela 5.4).

A RLM foi realizada através do *Software STATIGRAPHICS Plus 5.1*. Nos testes com esta ferramenta, os conjuntos de dados utilizados (treinamento e teste) foram os mesmos das RNAs, a fim de obter parâmetros de comparação entre as duas ferramentas matemáticas. O erro para a saída da RLM, obtida para a 1ª aplicação, foi de 0,011215 e para a 2ª aplicação foi de 0,0879.

Embora ambas as técnicas tenham apresentado resultados satisfatórios, as RNAs apresentaram um desempenho superior quando comparado ao método estatístico. Para a 1ª aplicação, tem-se os valores dos erros 0,00544 (RNAs) e 0,011215 (RLM), ou seja, apesar de ambos serem muito pequenos, o erro encontrado na RNA é praticamente a metade do valor do erro encontrado com a técnica de RLM. Da mesma forma, tem-se os valores para a 2ª aplicação, tem-se os valores dos erros 0,07108 (RNAs) e 0,0879 (RLM).

Assim, a melhor maneira de prever a duração do tempo de audiências trabalhistas, é utilizar-se da RNA “E30N10I30”, da tabela 5.2, onde os pesos foram gerados através do 2º teste (atributos codificados com ACP). Desta forma, dado um novo processo para o qual se deseja fazer o agendamento “inteligente” da audiência trabalhista deve-se, inicialmente, codificar os seus atributos e determinar as suas Componentes Principais e, em seguida, utilizando a topologia

## PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

e os pesos da rede “E30N10I30”, obter o número de minutos necessários para o mesmo. Desta forma cada audiência a ser agendada junto à justiça do trabalho do fórum de SJP, PR, apresentará um tempo estimado de duração, possibilitando que a agenda das audiências deixe de apresentar intervalos de tempo iguais para todos os casos, passando a disponibilizar um tempo específico para cada caso.

Já para prever o tempo de duração do trâmite de um novo processo trabalhista, deve-se utilizar a RNA “E23N6I50”, da tabela 5.4, onde os pesos foram gerados através do 3º teste (atributos ordinais, sem codificação, e com ACP). Assim, dado um novo processo trabalhista para o qual se deseja conhecer o tempo de duração do trâmite, deve-se determinar as Componentes Principais deste processo e, em seguida, utilizando a topologia e os pesos da rede E23N6I50, obter o número de meses necessários para o mesmo.

Vale salientar que de tempos em tempos, de acordo com a sugestão dos especialistas da área (juízes do trabalho), dados (processos) mais recentes, com respostas conhecidas e confiáveis, deverão ser incluídos na base de dados e a metodologia deverá ser repetida, sempre vislumbrando o menor erro possível para as referidas previsões. Com isto, espera-se obter um maior dinamismo e acurácia do sistema judiciário, assim como uma maior satisfação dos seus usuários.

### Referências

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

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

**Baesens, B.; Setiono, R.; Mues, C. & Vanthienen, J. (2003).** Using Neural Network Rule Extraction and Decision Tables for Credit-Risk Evaluation. *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.** Neuro-CPT: Classificação de Solos usando-se Redes Neurais Artificiais. *Engevista*, v. 8, p. 37-48, 2006.

**Bond, M.T.; Seiler, V. L.; Seiler, M. J. (2002).** Residential 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.** Análise Econômico-Financeira de Empresa Sob a Ótica da Estatística Multivariada. Curitiba, 2002. Dissertação de Mestrado, PPGMNE, UFPR.

**Lu, H.; Setiono, R. & Liu, H., (1996).** Effective Data Mining Using Neural Networks. *IEE Transactions on Knowledge and 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.

**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.



## PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

**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.; 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.** Regimento Interno. Disponível em: <<http://www.trt9.gov.br/>> Acesso em 07 outubro 2007.

**TST - Tribunal Superior do Trabalho.** Atribuições. Disponível em: <<http://www.tst.gov.br/>> Acesso em 16 fevereiro 2007.

ANEXO

Tabela 1. Atributos considerados no trabalho e as suas respectivas codificações tornando-os binários.

Atributos	Valores originais dos Atributos	Intervalos	Entrada 1	Entrada 2	Entrada 3	Entrada 4	Número de Entradas
1. Tipo de Rito (atributo nominal)	1 = RT	1 = RT	1	---	---	---	1
	0 = PS	0 = PS	0	---	---	---	
2. Tempo de Serviço (atributo ordinal)	<b>Faixa</b>	<b>meses</b>					3 4
	1	>=1 e < 15	0	0	0	---	
	2	>=15 e < 35	0	0	1	---	
	3	>=35 e < 85	0	1	1	---	
	4	>=85	1	1	1	---	
3. Salário do Reclamante (atributo ordinal)	<b>Faixas</b>	<b>Salário em reais</b>					4 3
	1	>= 330 e < 450	0	0	0	0	
	2	>=450 e < 620	0	0	0	1	
	3	>=620 e < 800	0	0	1	1	
	4	>=800 e < 1340	0	1	1	1	
	5	>=1340	1	1	1	1	
4. Profissão (atributo nominal)	<b>Setor</b>	<b>Cargo</b>					4 4
	Comércio	Direção	0	0	1	1	
	Comércio	Execução	0	0	1	0	
	Indústria	Direção	0	1	0	1	
	Indústria	Execução	0	1	0	0	
	Serviço	Direção	1	0	0	1	
5.1 Objeto do Processo – Falta de Registro em Carteira (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0 = não	0	---	---	---	
5.2 Objeto do Processo – Diferenças Salariais (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0 = não	0	---	---	---	
5.3 Objeto do Processo – Verbas Rescisórias (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0 = não	0	---	---	---	
5.4 Objeto do Processo – Multa Art 477 (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0 = não	0	---	---	---	
5.5 Objeto do Processo – Multa Art 467 (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0 = não	0	---	---	---	
5.6 Objeto do Processo – Horas Extras (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0 = não	0	---	---	---	
5.7 Objeto do Processo – FGTS (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0 = não	0	---	---	---	
5.8 Objeto do Processo – Danos Morais (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0 = não	0	---	---	---	
5.9 Objeto do Processo – Seguro Desemprego (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0 = não	0	---	---	---	
5.10 Objeto do Processo – Vale Transporte (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0 = não	0	---	---	---	

(continua...)

## PESQUISA OPERACIONAL PARA O DESENVOLVIMENTO

(...continuação)

5.11 Objeto do Processo – Adicional de Insalubridade (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0= não	0	---	---	---	
5.12 Objeto do Processo – Adicional Noturno (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0= não	0	---	---	---	
**5.13 Objeto do Processo – Plano de Saúde (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0= não	0	---	---	---	
**6. Juiz (atributo nominal)	<b>Juiz</b>	---					3
	1		1	0	0		
	2		0	1	0		
	3		0	0	1		
**7. Depoimento (atributo ordinal)	No. de depoimentos	---	---				1
8. Acordo (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0= não	0	---	---	---	
*9. Audiências (atributo ordinal)		No. audiências					4
	1	1	0	0	0	0	
	2	2	0	0	0	1	
	3	3	0	0	1	1	
	4	4	0	1	1	1	
*10. Perícia (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0= não	0	---	---	---	
*11. Recurso Ordinário (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
	0 = não	0= não	0	---	---	---	
*12. Recurso Revista (atributo nominal)	1 = sim	1 = sim	1	---	---	---	1
<b>TOTAL: 12</b>							<b>37</b>

\*os atributos 9, 10, 11 e 12 não são considerados para a previsão do tempo de audiência (total de atributos codificados = 30 (37 – 7));

\*\*os atributos 5.13, 6 e 7 não são considerados para a previsão do tempo de trâmite (total de atributos codificados = 32 (37 – 5)).