

**REDES NEURAIS ARTIFICIAIS NA PREVISÃO DO TEMPO DE
DURAÇÃO DE AUDIÊNCIAS TRABALHISTAS****Alessandra Memari Pavanelli¹****Cap QCO Genival Pavanelli²**

RESUMO: O presente trabalho tem como objetivo prever o tempo de duração de audiências trabalhistas, otimizando dessa maneira, o agendamento dos horários dos juízes do Fórum Trabalhista de São José dos Pinhais (SJP), Paraná. A metodologia proposta neste trabalho consiste em comparar duas ferramentas: Rede Neural Artificial (RNA) e Regressão Linear Múltipla. Para tanto, utilizou-se os dados dos processos do Fórum Trabalhista de SJP. Foram treinadas diversas RNAs com várias topologias e, em alguns casos, utilizando a Análise de Componentes Principais, buscando melhorar ainda mais os resultados obtidos. Na Regressão Linear Múltipla os testes foram realizados com os mesmos conjuntos de dados utilizados nas RNAs a fim de comparar as duas técnicas. As ferramentas utilizadas mostraram-se bem eficientes, sendo que para o problema abordado, o desempenho das RNAs, com a utilização de Análise de Componentes Principais, foi superior ao resultado apresentado pela técnica de Regressão Linear Múltipla.

Palavras-chave: Audiências Trabalhistas; Rede Neural Artificial; Regressão Linear Múltipla.

1 - Professora de Física no Colégio Militar de Curitiba
Mestre em Métodos Numéricos em Engenharia – PPGMNE – UFPR
E-mail: alessandracmc@bol.com.br

2 - Professor de Física no Colégio Militar de Curitiba
Mestre em Métodos Numéricos em Engenharia – PPGMNE – UFPR
E-mail: pavanellimc@bol.com.br

INTRODUÇÃO

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) e cerca de 270 novas Varas do Trabalho foram criadas, a partir de 2003, 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, 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. Em virtude do aumento do número de processos trabalhistas, como consequência da industrialização maciça no município, faz-se necessária a expansão e a agilidade nos serviços da justiça. Com isso, a necessidade do uso de ferramentas matemáticas para o agendamento inteligente das audiências é de fundamental importância para essa agilidade na justiça.

O objetivo deste trabalho é estimar o tempo aproximado de cada audiência do Fórum Trabalhista de SJP, a fim de obter um agendamento “inteligente” dessas audiências. Para tanto, utilizar-se-á de dois métodos: um na área de inteligência artificial, Redes Neurais Artificiais e outro estatístico, Regressão Linear Múltipla. O objetivo de utilizar os dois métodos é o de se fazer uma comparação entre os resultados finais obtidos e, com isso, verificar qual fornece o melhor resultado.

O presente trabalho está estruturado da seguinte forma: na Seção 2 está a descrição do problema, levantamento e tratamento dos dados. Na Seção 3 está a metodologia do trabalho, momento em que são apresentados conceitos que envolvem as técnicas de RNAs, Análise de Componentes Principais e Regressão Linear Múltipla. A Seção 4 descreve a implementação computacional das técnicas utilizadas e a análise dos resultados. Finalizando, na Seção 5 são apresentadas as conclusões obtidas através da análise dos resultados da seção anterior.

DESCRIÇÃO DO PROBLEMA, LEVANTAMENTO E TRATAMENTO DE DADOS

Os dados utilizados neste trabalho foram obtidos junto ao Fórum Trabalhista de SJP, que hoje conta com a 1ª e 2ª Varas, trabalhando em ritmo acelerado e, em virtude dessa grande demanda de processos trabalhistas, está prestes a inaugurar a 3ª Vara. Somente no ano de 2006, o Fórum Trabalhista recebeu cerca de seis mil processos, número este que vem aumentando ano após ano.

Com o propósito pré-estabelecido de treinar uma RNA e também a aplicação da técnica da Regressão Linear Múltipla, a fim de obter a previsão da duração de uma audiência trabalhista, foram realizadas várias reuniões com o juiz titular do Fórum Trabalhista de SJP, com o objetivo de determinar quais dados seriam relevantes. Tais dados, também chamados de atributos da RNA, são citados a seguir.

- a. Rito: pode ser rito trabalhista (RT) ou procedimento sumaríssimo (PS);
- b. Tempo de Serviço: é dado em meses pela diferença entre a data de dispensa e data de admissão;
- c. Salário do Reclamante: último salário recebido;
- d. Profissão: função exercida pelo reclamante. Dividiu-se este atributo em duas partes: setor, que 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. São elas: 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. Tempo da Audiência: é dado em minutos pela diferença entre o término e o início da audiência.

Os processos analisados foram ajuizados no ano de 2006, na 1ª Vara do Trabalho do Fórum de SJP. Foram coletados dados de 108 processos através de uma Ficha de Dados. Os atributos listados anteriormente serviram como dados para criar a matriz destinada ao treinamento e ao teste da RNA bem como para aplicação da técnica de Regressão Linear Múltipla.

A maioria dos dados foi tratada de forma a corresponder a uma ou mais coordenadas binárias (LU et al, 1996), (BAESENS et al, 2003) do vetor de entrada para as técnicas utilizadas, conforme comentado a seguir:

- quando o atributo “Rito” for do tipo “RT”, apresentará entrada com valor “1”; quando for do tipo “PS”, apresentará valor “0”, caracterizando-a como uma variável nominal;
- o atributo “Tempo de Serviço” foi dividido em faixas, mantendo-se a mesma cardinalidade dentro de cada uma delas. Devido ao fato de ser uma variável ordinal, as coordenadas binárias que a representam foram designadas conforme a Tabela 1 a seguir, apresentando assim três coordenadas no vetor de entrada.

TABELA 1 – Codificação do Atributo “Tempo de Serviço”

Faixa	Duração	Entrada
1	≥ 1 e < 15 meses	0 0 0
2	≥ 15 e < 35 meses	0 0 1
3	≥ 35 e < 85 meses	0 1 1
4	≥ 85	1 1 1

- o atributo “Salário”, da mesma maneira que o atributo “Tempo de Serviço”, foi dividido em faixas procurando manter a mesma cardinalidade dentro de cada uma delas. Conforme mostra a Tabela 2, a seguir, este atributo fornece quatro coordenadas para o vetor de entrada;

TABELA 2 – Codificação do atributo “Salário do Reclamante”

Faixa	ÚLTIMO SALÁRIO	Entrada
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

- o atributo “profissão”, classificado como variável nominal, foi dividido em “setor” e “cargo”, conforme citado anteriormente. De acordo com a Tabela 3, a seguir, este atributo apresentará quatro coordenadas no vetor de entrada;

TABELA 3 – Codificação do atributo “Profissão”

Setor	Cargo	Coordenadas		Entrada
		Do Setor	Do Cargo	
Comércio	Direção	0 0 1	1	0 0 1 1
	Execução	0 0 1	0	0 0 1 0
Indústria	Direção	0 1 0	1	0 1 0 1
	Execução	0 1 0	0	0 1 0 0
Serviço	Direção	1 0 0	1	1 0 0 1
	Execução	1 0 0	0	1 0 0 0

- o atributo “Objeto do Processo” possui 13 coordenadas binárias no vetor de entrada. Cada uma das coordenadas corresponde a uma das solicitações do autor do processo, conforme já mencionado. Apresentará valor “1”, se for solicitada; valor “0”, em caso de não solicitação;

- como foram analisados processos de três juizes diferentes, e por se tratar de um atributo nominal, cada um deles foi codificado conforme a Tabela 4, a seguir;

TABELA 4 – CODIFICAÇÃO DO ATRIBUTO “JUIZ”

Juiz	Entrada
1	1 0 0
2	0 1 0
3	0 0 1

- como o número de depoimentos em cada audiência não era superior a quatro, manteve-se este atributo com o seu valor original;
- quando houver “Acordo” entre as partes antes do julgamento, este atributo apresentará o valor “1”, caso contrário apresentará o valor “0”.

Sendo assim, o vetor de entrada dos dados possui 30 coordenadas, que são os atributos codificados do processo conforme especificado anteriormente. A matriz de entrada de dados possui então 108 linhas e 30 colunas, considerando-se que os dados foram retirados dos 108 processos (padrões).

Podem-se citar algumas publicações relacionadas a este trabalho: o trabalho de PAVANELLI (2007) tem como objetivo fornecer uma previsão do tempo de duração de um processo trabalhista. Os dados também foram obtidos junto à 1ª Vara do Trabalho de SJP. Foram analisados 100 processos, dos quais 75 serviram para o treinamento da RNA e os outros 25 serviram de testes. A técnica utilizada foi RNA com o algoritmo back-propagation e função de ativação sigmoidal, que fornece saídas contidas num intervalo entre “0” e “1” para entradas contidas num intervalo de $(-\infty, \infty)$. Foi realizado o treinamento supervisionado, e diversos “testes”, com o grupo de treinamento variando o número de neurônios na camada escondida, bem como o número de iterações, de forma a obter o melhor resultado com o menor erro global no grupo de testes. Em alguns “testes”, utilizou-se a Análise de Componentes Principais, buscando melhorar ainda mais os resultados obtidos. A ferramenta utilizada apresentou-se bem eficiente e forneceu bons resultados, com uma margem de erro dentro do esperado.

BAPTISTELLA et al (2006) procuram técnicas alternativas com o objetivo de determinar valores venais para imóveis de Guarapuava – PR, pois é sabido que não existe nenhuma técnica para calcular o valor de um imóvel, apenas a experiência. O método utilizado é a comparação de técnicas matemáticas, no caso, é utilizado RNA com múltiplas camadas com diversas topologias e Análise de Regressão Linear Múltipla. Os dados foram obtidos de um Cadastro Imobiliário da cidade, fornecido pelo Setor de Planejamento da Prefeitura Municipal. As duas técnicas se mostraram eficientes, porém a RNA forneceu um resultado superior.

ADAMOWICZ (2000) utiliza técnicas de reconhecimento de padrões tais como RNA 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. As duas técnicas se apresentaram eficientes na discriminação das empresas, porém o desempenho da RNA 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 através de RNA do tipo perceptron multicamadas, que funcionou como um classificador de padrões. O algoritmo de treinamento foi o back-propagation, e a função de ativação foi do tipo sigmoidal. 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 a RNA cumpre bem seu papel como classificador de padrões.

SOUSA et al (2003) utilizam técnicas de RNA com três camadas de neurônios com o algoritmo back-propagation. O objetivo é 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 da RNA às amostras comerciais foi inadequada devido à diferença de ingredientes das salsichas usadas no treinamento e os ingredientes das amostras comerciais.

BIONDI NETO et al (2007) mostram em seu trabalho que, atualmente, a determinação do tipo de solo é obtida através de ábacos; o objetivo dessa pesquisa é aplicar um método computacional para classificar o solo. A técnica utilizada é RNA com método 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. Os resultados apresentaram um acerto de 95% na classificação dos solos.

METODOLOGIA DO TRABALHO

A metodologia aplicada neste trabalho buscou, por meio da utilização das RNAs e da Regressão Linear Múltipla, reconhecer padrões dentro de um processo trabalhista, a fim de fazer uma estimativa do tempo de duração de cada audiência trabalhista. Pode-se, de maneira sucinta, mostrar o desenvolvimento das RNAs nos últimos 60 anos, como se segue (STEINER, 1995) e também um breve histórico da Regressão linear Múltipla (LIMA,2002):

Data de 1943, com McCulloch e Pitts, o primeiro modelo de neurônio artificial cuja coleção era capaz de calcular algumas funções lógicas. Em 1949, Hebb desenvolveu uma regra de aprendizagem básica. Em 1959, Rosenblatt criou o Perceptron, primeiro modelo de rede neural. Em 1962, Widrow desenvolveu o Adaline, um processador que possuía uma poderosa estratégia de aprendizado. Em 1969, o estudo das RNAs foi praticamente abandonado devido à repercussão do trabalho de Minsky e Papert, que expunha as limitações do Perceptron. Em 1974, Werbos lançou as bases do algoritmo back-propagation, que apresentava grande capacidade de aprendizado. A

partir de meados da década de 80, consagrou-se a técnica de treinamento por retro-propagação, fato que gerou uma explosão de interesse pelas RNAs na comunidade internacional (HAYKIN, 2001).

Apresenta-se a seguir, na Figura 1, um modelo de RNA.

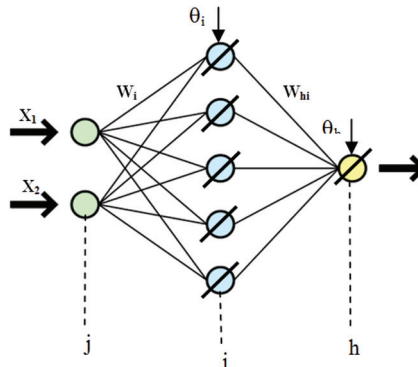


FIGURA 1. Exemplo de RNA com duas Camadas.

A RNA implementada neste trabalho é classificada como Rede de Múltiplas Camadas ou Rede feed-forward (alimentada para frente) e 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 $+\infty$.

Para verificar a performance da rede foi utilizado o MSE (erro médio quadrático).

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

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 para o padrão p (pela rede).

Apresenta-se na Figura 2, um esquema do algoritmo implementado quando $h = 1$, ou seja, quando a rede apresenta um único neurônio na saída.

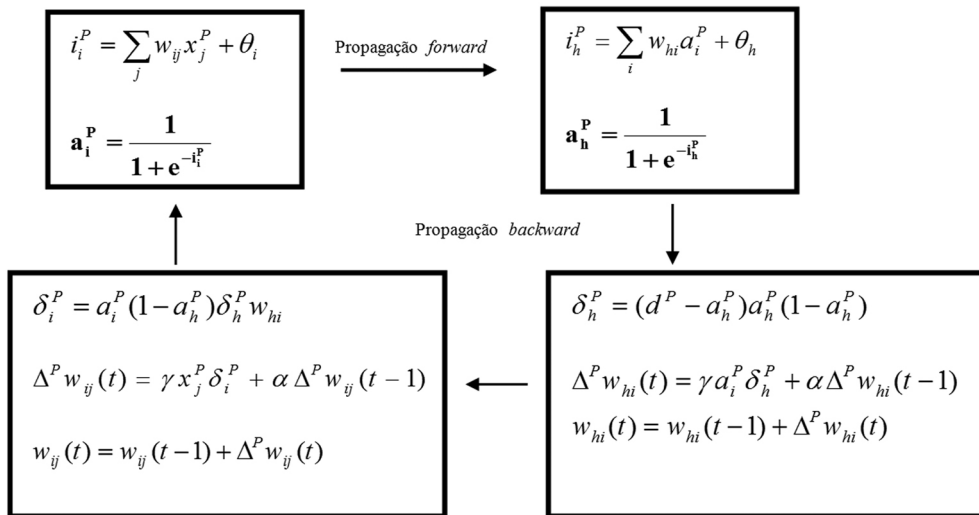


FIGURA 2. Ilustração das propagações *forward* (para frente) e *backward* (para trás) do algoritmo *back-propagation*:

Nesta Figura 2, tem-se:

p = padrão; t = iteração; γ = taxa de aprendizagem; α = taxa de momento;

w_{ij} e w_{hi} = pesos entre as camadas de entrada-oculta e oculta-saída, respectivamente;

i^p = potencial interno do processador; x_j^p = *input* para o padrão p ; a^p = *output* para o padrão p ; d^p = valor desejado para o padrão p .

O segundo método utilizado neste trabalho 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 utilizados em diversas áreas do conhecimento.

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 binárias (dicotômicas), ou seja, respostas “0” ou “1”.

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 .

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; β é o vetor de parâmetros a serem estimados; ε vetor de erros aleatórios.

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

IMPLEMENTAÇÃO COMPUTACIONAL E OBTENÇÃO DE RESULTADOS

Conforme descrito na Seção 2, os métodos propostos neste trabalho foram aplicados após a coleta e o tratamento dos dados dos 108 processos analisados, que foram ajuizados no ano de 2006, junto ao Fórum Trabalhista de SJP. Todos os dados obtidos de cada um dos processos trabalhistas serviram para compor a matriz de entrada. O treinamento da RNA 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 da RNA foi implementado um programa utilizando o Software Visual Basic 6.0.

Para a realização do treinamento da RNA foi utilizado o algoritmo supervisionado back-propagation e a função de ativação, em todos os neurônios, foi a sigmoideal, como já mencionado, que fornece saídas entre (0, 1) para entradas no intervalo $(-\infty, \infty)$. Devido a essas condições da função de ativação, foi necessário enquadrar as saídas, ou seja, os tempos de duração de cada audiência, que variam de 1 a 93 minutos, foram divididos por 93 enquadrando-os ao intervalo de abrangência da função.

Do conjunto de 108 processos cadastrados, 75% foram utilizados para o treinamento da rede e os 25% restantes foram utilizados no teste. Foram utilizados, em todos os testes, quatro conjuntos de pesos iniciais.

Foram realizados três tipos de testes. No primeiro a maioria das entradas foi codificada, conforme mencionado na Seção 2. No segundo teste, além das entradas serem codificadas, também foi realizada a Análise de Componentes Principais. No terceiro as variáveis ordinais, ou seja, os atributos salário, tempo de serviço e número de audiências, não foram codificados e a matriz de dados foi submetida ao processo de Análise de Componentes Principais a fim de avaliar a importância relativa das variáveis que a compõem.

Em todos os testes realizados com 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 menor MSE no grupo de teste. A melhor arquitetura obtida voltou a ser treinada, agora variando-se o número de iterações, sempre buscando o menor erro no grupo de teste.

Foi escolhida uma nomenclatura para cada topologia a fim de representar na seqüência as seguintes características: quantidade de entradas, número de neurônios na camada escondida e número de iterações. Por exemplo, a rede E32N1I40 é uma rede com 32 entradas, 1 neurônio na camada escondida e foi treinada com 40 iterações.

Dentre os testes realizados o que apresentou melhor resultado, ou seja, aquele que apresentou menor erro foi o Teste 2 onde os dados foram codificados conforme descrito na Seção 2, em seguida a matriz de entrada foi submetida à Análise de Componentes Principais. A Tabela 5. a seguir mostra a variação do número de neurônios da RNA deste teste.

TABELA 5 – RESULTADOS DAS SIMULAÇÕES COM VARIAÇÃO DO NÚMERO DE NEURÔNIOS

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	10	E30N10I50	0,00046	0,00562
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				

Verifica-se a partir da Tabela 5 que a melhor topologia foi a E30N10I50, ou seja, a RNA com 30 neurônios na camada de entrada, 10 neurônios na

camada escondida e foi treinada com 50 iterações. O MSE foi igual a 0,00046 para o grupo de treinamento e 0,00562 para o grupo de teste.

A RNA 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 6 a seguir.

TABELA 6 – RESULTADOS DAS SIMULAÇÕES COM VARIAÇÃO DO 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
3	E30N10I30	0,00090	0,00544	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 na Tabela 6 é obtido na terceira simulação, cuja topologia é E30N10I30 onde se tem 30 iterações e MSE igual a 0,00544 no grupo de teste.

Na aplicação da técnica de Regressão Linear utilizou-se o mesmo conjunto de dados das RNAs. As tabelas de dados foram construídas no Software Excel, e para a validação dos resultados do método de Regressão Linear Múltipla foi utilizado o 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. No primeiro teste, as entradas foram codificadas sem a ACP. No segundo teste as entradas foram codificadas e foi realizado ACP, e no terceiro os dados não foram codificados e foi aplicado a ACP.

O melhor resultado obtido neste método foi também o Teste 2 em que os dados foram codificados conforme descrito na Seção 2, em seguida, a matriz de entrada foi submetida à ACP. A equação de Regressão Linear Múltipla que descreve a relação entre a variável resposta (tempo de audiência) e as variáveis independentes nesse teste é descrita a seguir:

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

Ao aplicar a equação acima nos 108 processos, obtém-se um erro quadrático médio igual a 0,011215.

CONCLUSÕES

O Fórum Trabalhista de SJP vem aumentando de maneira considerável o número de processos trabalhistas. Diante desse aumento, faz-se necessária a utilização de ferramentas matemáticas que permitam uma previsão de duração do tempo de audiências trabalhistas.

O presente trabalho teve como objetivo comparar as técnicas de RNAs e Regressão Linear Múltipla, para encontrar a melhor previsão de duração do tempo das audiências trabalhistas.

Utilizando os dados obtidos de processos do Fórum Trabalhista de SJP, que são as entradas para ambas as técnicas, as mesmas foram treinadas buscando obter de forma automática uma previsão de duração das audiências.

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 codificação dos atributos de entrada, o número de neurônios na camada escondida, o conjunto de pesos iniciais e o número de iterações, fornecendo como resposta (saída da RNA), a duração do tempo das audiências de cada um dos processos trabalhistas.

A Regressão Linear Múltipla foi realizada através do Software STATIGRAPHICS Plus 5.1. Nos testes com esta ferramenta, os conjuntos de dados utilizados foram os mesmos das RNAs, a fim de obter parâmetros de comparação entre as duas ferramentas matemáticas.

Na comparação das duas técnicas, embora ambas tenham apresentado resultados satisfatórios, as RNAs apresentaram um desempenho superior quando comparado ao método estatístico. O menor erro para a RNA tem o valor de 0,00544 e o da Regressão Linear Múltipla igual a 0,011215, 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 Regressão Linear Múltipla.

Assim, a melhor maneira de prever a duração do tempo de audiências trabalhistas é utilizar-se da RNA com pesos gerados no Teste 2, utilizando-se todas as componentes principais. Assim, dado um novo processo trabalhista para o qual se deseja fazer o agendamento “inteligente”, pode-se, utilizando os pesos do referido Teste 2, obter-se o número de minutos necessários para o mesmo obtendo-se, assim, um maior dinamismo do sistema judiciário. Vale salientar que, de tempos em tempos, o treinamento da RNA, com a Análise das Componentes Principais, deve ser refeito, atualizando-se com isso as informações da rede com dados recentes.

REFERÊNCIAS

ADAMOWICZ, E. C. **Reconhecimento de Padrões na Análise Econômico-Financeira de Empresas**. Curitiba, 2000. Dissertação de Mestrado - Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

AMBRÓSIO, P. E. **Redes Neurais Artificiais no Apoio ao Diagnóstico Diferencial de Lesões Intersticiais Pulmonares**. Ribeirão Preto, 2002. Dissertação de Mestrado - Universidade de São Paulo.

BAESENS, B.; SETIONO, R.; MUES, C. & VANTHIENEN, J. (2003), **Using Neural Network Rule Extraction and Decision Tables for Credit-Risk Evaluation**. *Management Science Informs*, vol. 49, n° 3, p. 312-329.

BAPTISTELA, M.; STEINER, M. T. A.; CHAVES NETO, A. (2006), **O Uso de Redes Neurais e Regressão Linear Múltipla na Engenharia de Avaliações**: Determinação dos Valores Venais de Imóveis Urbanos, Atas do XXXVIII SBPO.

BIONDI NETO, L.; SIEIRA, A. C. C. F.; DANZIGER, B. R.; SILVA, J. G. S.(2007), **Classificação de Solos Usando-se Redes Neurais Artificiais**, Atas do XXXIX SBPO.

HAYKIN, S., **Redes Neurais: Princípios e Prática**, Bookman, 2001.

LIMA, J. D., **Análise Econômico-Financeira de Empresa Sob a Ótica da Estatística Multivariada**. Curitiba, 2002. Dissertação de Mestrado - Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

LU, H.; SETIONO, R. & LIU, H., (1996). **Effective Data Mining Using Neural Networks**. *IEE Transactions on Knowledge and Data Engineering*, vol. 8, n° 6, p.957-961.

PAVANELLI, G., **Análise do Tempo de Duração de Processos Trabalhistas Utilizando Redes Neurais Artificiais Como Apoio à Tomada de Decisões**. Curitiba, 2007. Dissertação de Mestrado - Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

SOUSA, E. A.; TEIXEIRA, L. C. V.; MELLO, M. R. P. A.; TORRES, E. A. F. S.; MOITA NETO, J. M. **Aplicação de Redes Neurais para Avaliação do Teor de Carne Mecanicamente Separada em Salsicha de Frango**. *Ciência e Tecnologia de Alimentos*, vol.23 no.3 Campinas Sept./Dec. 2003.

STEINER, M. T. A. Uma Metodologia Para o Reconhecimento de Padrões Multivariados com Resposta Dicotômica. Florianópolis, 1995. Tese de Doutorado – Programa de Pós Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina.

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.