

Redes Neurais Artificiais, Regressão Linear Múltipla e Árvores de Decisão Aplicadas à Previsão junto à Justiça Trabalhista

Genival Pavanelli

Universidade Federal do Paraná – UFPR
Programa de Pós-Graduação em Métodos Numéricos em Engenharia
CP19081 – Curitiba, PR; CEP 81531-990
pavanellimc@bol.com.br

Alessandra Memari Pavanelli

Universidade Federal do Paraná – UFPR
Programa de Pós-Graduação em Métodos Numéricos em Engenharia
CP19081 – Curitiba, PR; CEP 81531-990
alessandracmc@bol.com.br

Deise Maria Bertholdi Costa

Universidade Federal do Paraná – UFPR
Departamento de Expressão Gráfica - Programa de Pós-Graduação em Métodos Numéricos em
Engenharia
CP19081 – Curitiba, PR; CEP 81531-990
deise@ufpr.br

RESUMO

O presente trabalho tem como objetivo prever o tempo de duração do trâmite de processos trabalhistas. A metodologia proposta consiste em aplicar e comparar três técnicas da área da Programação Matemática, Redes Neurais Artificiais (RNAs), Regressão Linear Múltipla (RLM) e Árvores de Decisão com o intuito de obter o melhor desempenho possível para a previsão. Para tanto, utilizou-se os dados de processos trabalhistas do Fórum Trabalhista de São José dos Pinhais, Paraná, Brasil, para se fazer o treinamento de diversas RNAs, da RLM e da Árvore de Decisão. Assim, tendo-se novos dados (processos) para os quais seja necessário prever o tempo de duração do processo, tenha-se condições de “diagnosticar” esse tempo, preliminarmente ao seu trâmite. As três técnicas utilizadas mostraram-se eficientes, apresentando resultados consistentes com uma margem de erro aceitável.

PALAVRAS CHAVE: Previsão, Processos Trabalhistas, Programação Matemática.

ABSTRACT

This paper aims to predict the duration of labor proceedings. 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. 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.

KEYWORDS: Forecasting, Labor Proceedings, Mathematical Programming.

1. Introdução

Este trabalho apresenta uma proposta de aplicação de técnicas da área da Pesquisa Operacional, junto à justiça do trabalho. Tal proposta consiste em fornecer uma previsão do tempo de duração de um processo trabalhista aos usuários do Fórum Trabalhista de São José do Pinhais, PR, Brasil.

A fim de obter tal previsão, utilizou-se três métodos: um da área de inteligência artificial, Redes Neurais Artificiais (RNAs) e dois, da área estatística, Regressão Linear Múltipla (RLM) e Árvore de Decisão. O objetivo de utilizar estes três 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), RLM e Árvore de Decisão. 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.

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 *Bene 2* (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 Mota e Steiner (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. 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.

Souza *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 Syange (2009), é proposta a utilização de um modelo de Programação Linear (PL) 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 segundo modelo matemático que faz uso do primeiro, de forma a garantir a obtenção de bobinas de boa qualidade a um mínimo custo.

Biondi Neto *et al.* (2006) mostram em seu trabalho a determinação do tipo de solo, que, até então, só 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 (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 verificada, 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.

Steiner *et al.* (2007) utilizam técnicas de extração de regras tais como o *Neurorule* e o *software* WEKA para extrair regras de uma rede neural artificial treinada. A RNA classifica as empresas como boas ou más tomadoras de crédito. A partir da rede treinada realizaram três tipos de testes de extração de regras. No primeiro a extração de regras foi feita diretamente dos dados originais, já no segundo teste descartou-se os padrões classificados erroneamente pela RN, enquanto que no terceiro teste além de descartar os padrões classificados erroneamente pela RN codificou-se os atributos segundo as codificações "termômetro" e "*dummy*", tornando-os binários. Os resultados foram bastante satisfatórios apresentando acurácia superior a 80% quanto a concessão (ou não) de crédito bancário.

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

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 a previsão do tempo de duração do trâmite dos processos trabalhistas, é de fundamental importância para essa agilização.

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, Brasil. Com o objetivo de determinar quais atributos seriam relevantes 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 10 atributos listados 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. Acordo: quando ocorre o acordo entre as partes;
- g. Perícia: se há necessidade ou não da realização de alguma espécie de perícia como, por exemplo, a perícia médica ou a de periculosidade;
- h. Recurso Ordinário: quando uma das partes (reclamante ou reclamado) não concorda com a sentença emitida pelo juiz e solicita Recurso Ordinário ao TRT;
- i. 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;
- j. Número de Audiências: trata do número de audiências necessárias para que o juiz emita a sentença;

Os 10 atributos listados anteriormente, utilizados para a previsão do tempo de duração do processo foram coletados de 100 processos dando origem a matriz destinada ao treinamento e ao teste das RNAs, bem como para aplicação da técnica de RLM e construção da Árvore de Decisão.

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, cada um dos 10 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 Categórica	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

A partir da codificação acima explicitada os 10 atributos passaram a fornecer 32 entradas para as RNAs; logo, a matriz de conta com 100 linhas e 32 colunas.

4. Metodologia do Trabalho

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

Buscando minimizar o erro das técnicas aplicadas foram realizados três testes distintos. No primeiro teste todos os atributos foram codificados, conforme descrito na seção 3.1, de forma que cada padrão apresentasse um vetor de entrada com 32 coordenadas binárias. No segundo teste a matriz de dados codificada (conforme teste anterior) foi submetida a ACP, visando avaliar a importância relativa das variáveis que compõem a amostra de dados. No terceiro teste realizado as variáveis ordinais originais não foram codificadas, ou seja, os atributos salário, tempo de serviço e número de audiências não foram transformados em vetores binários e em seguida a matriz foi submetida a ACP, de tal forma que cada padrão apresentasse, para este teste, um vetor de entrada com 23 coordenadas.

4.1 Redes Neurais Artificiais

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).

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

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 .

4.2 Regressão Linear Múltipla

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 , 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: $Y = X\beta + \mathcal{E}$, onde Y é a variável resposta; X matriz do modelo; β é o vetor de parâmetros a serem estimados; \mathcal{E} vetor de erros aleatórios.

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

4.3 Árvores de Decisão

As árvores de decisão constituem uma técnica muito poderosa, amplamente utilizada, baseada numa hierarquia de testes a algumas das variáveis envolvidas em um problema de decisão. O conhecimento adquirido a partir dessa técnica é expresso através de regras, fato este que justifica sua ampla utilização. Pode ser utilizada com duas finalidades: previsão (exemplo: descobrir se um cliente será um bom pagador em função de suas características) e descrição (fornecer informações interessantes a respeito das relações entre os atributos preditivos e o atributo classe numa base de dados).

Sua estrutura possui as seguintes características:

- cada nó interno é um teste em um atributo preditivo;
- uma ramificação partindo de um nó interno representa um resultado para o teste;
- uma folha da árvore representa um rótulo de classe.

Para classificar um exemplo desconhecido basta encaminhá-lo para baixo na árvore de acordo com os valores dos atributos testados em nós sucessivos, e, quando uma folha é atingida a instância é classificada de acordo com a classe atribuída à folha (WITTEN e FRANK, 2005).

4.4 Análise de Componentes Principais

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, RLM e Árvores de Decisão, 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 das 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 trâmite dos processos, que variam de 1 a 94 meses, foram divididos por 94. Vale salientar que o tempo de duração do processo possui distribuição uniforme.

Para a avaliação da RNA, foi utilizado o procedimento *hold out*, ou seja, do conjunto dos processos cadastrados, 75% foram utilizados para o treinamento da rede e os 25% restantes foram utilizados no teste. Na aplicação das RNAs foram utilizados, em todos os testes, quatro conjuntos de pesos iniciais.

A Figura 5.1, a seguir, mostra formas conceituais de duas curvas de aprendizagem, uma em relação ao conjunto de treinamento e outra em relação ao conjunto de validação. As curvas são distintas, sendo que a curva de treinamento (aprendizagem) decresce monotonamente para

um número crescente de iterações. Já a curva de validação decresce até um mínimo e então volta a crescer à medida que o treinamento prossegue.

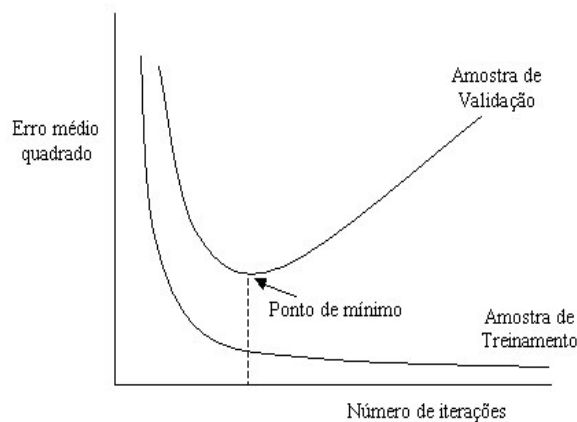


Figura 5.1. Treinamento Versus Capacidade de Generalização da RNA - Fonte: Haykin (2002)

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 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, conforme ilustrado na Figura 5.1.

Foi escolhida uma nomenclatura para cada topologia, a fim de representar, em sequência, 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 Obtidos

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 5.1, a seguir, mostra a variação do número de neurônios nas RNAs para este teste.

Tabela 5.1. 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
6	E23N6I50	0,02383	0,07108	13	E23N13I50	0,02485	0,11329
7	E23N7I50	0,02294	0,08402	14	E23N14I50	0,02186	0,21053

De acordo com a Tabela 5.1, 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 se variando o número de iterações, a fim de se obter o menor erro possível. A tabela 5.2, a seguir, mostra os resultados relevantes.

Tabela 5.2. 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	E23N6I50	0,02383	0,07108	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 5.2, 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 5.1 abaixo.

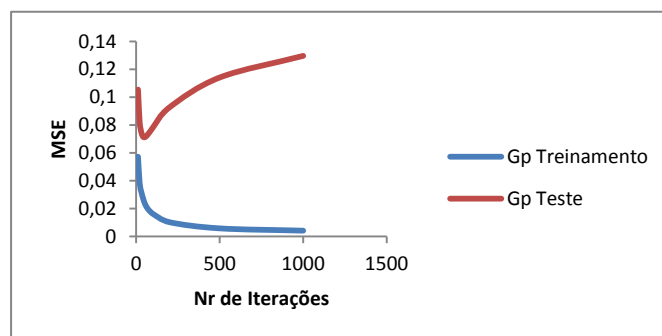


Gráfico 5.1. Erro do Grupo de Treinamento e do Grupo de Teste

A previsão, 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)

$$\begin{aligned} \text{Tempo de processo} = & 1.0e-003*(0,2422 - 0,0009*CP1 - 0,0001*CP2 - 0,0005*CP3 - \\ & 0,0005*CP4 - 0,0001*CP5 - 0,0002*CP6 + 0,0002*CP7 - 0,0001*CP8 + 0,0007*CP9 - \\ & 0,0003*CP10 - 0,0001*CP11 + 0,0001*CP12 + 0,0005*CP13 - 0,0002*CP14 - 0,0002*CP15 - \\ & 0,0003*CP16 + 0,0001*CP17 - 0,0003*CP18 + 0,0003*CP19 + 0,0003*CP20 - 0,0001*CP21 + \\ & 0,0001*CP22 + 0,0009*CP23) \end{aligned} \quad (5.2)$$

Ao se aplicar a equação de regressão aos mesmos conjuntos de treinamento e teste utilizados nas RNAs, obteve-se o MSE (conforme descrito na equação 4.1) igual a 0,0743 para o conjunto de treinamento, enquanto que no conjunto de teste o erro foi de 0,1287.

A técnica de Árvore de Decisão foi aplicada a partir do software WEKA (*Waikato Environment for Knowledge Analysis*), que é gratuito e de código aberto usado para minerar dados. Assim como na RLM, na aplicação da técnica Árvore de Decisão utilizou-se os mesmos conjuntos de dados e também realizou-se os três tipos de testes. Esta técnica também apresentou os melhores resultados no 3º teste, onde o erro quadrático médio apresentado foi de 0,0881.

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 objetivo de possibilitar a “negociação” entre as partes, através da previsão da duração do trâmite de processos trabalhistas.

A aplicação aqui apresentada, relacionada a processos da Justiça do Trabalho, mostra, mais uma vez, a grande aplicabilidade das técnicas da área de Pesquisa Operacional. A aplicação, aqui abordada, visa comparar as técnicas de RNAs, RLM e Árvores de Decisão para encontrar a melhor previsão. Com dados obtidos de 100 processos, que são as entradas para as técnicas, buscou-se obter, de forma automática, uma previsão de duração 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. A melhor resposta obtida apresentou um erro de 0,07108, para uma RNA com 23 neurônios na camada de entrada, seis neurônios na camada oculta, com 50 iterações (tabela 5.2).

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 obtido foi igual a 0,0743 para o conjunto de treinamento e de 0,1287 para o conjunto de teste.

A técnica Árvores de Decisão foi aplicada através do *software WEKA (Waikato Environment for Knowledge Analysis)*. Utilizaram-se os mesmos conjuntos de dados a fim de comparar as técnicas aplicadas. Nesta ferramenta o 3º teste também apresentou o melhor resultado, contando com um MSE total igual a 0,0881.

Embora as técnicas tenham apresentado resultados satisfatórios, as RNAs apresentaram um desempenho superior quando comparado aos outros métodos, conforme se pode observar pelos erros 0,07108 (RNAs), 0,0879 (RLM) e 0,0881 para Árvore de Decisão.

Assim, a melhor maneira de prever tempo de duração do trâmite de um novo processo trabalhista, é utilizar a RNA “E23N6I50”, da Tabela 5.2, 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 incluídas na base de dados e a metodologia deverá ser repetida, sempre vislumbrando o menor erro possível para a referida previsão. Com isto, espera-se obter um maior dinamismo e acurácia do sistema judiciário, assim como uma maior satisfação dos seus usuários.

Referências

- 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.
- 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. Residential Real Estate Prices: a Room with a View. The Journal of Real Estate Research, v. 23, n. 1, p. 129-137.
- Haykin, S.**, Redes Neurais: Princípios e Prática, Bookman, 2001.
- 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 and Data Engineering, vol. 8, n° 6, p.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.
- 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.** 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.
- Witten, I. H., Frank, E.**, 2005, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann Publishers, 2nd edition.