

O Uso de Redes Neurais na Engenharia de Avaliações: Determinação dos Valores Venais de Imóveis Urbanos

Neural Network Use in Evaluation Engineering: Figuring Out the Value of Urban Real Estate Tax

Marisa Baptistella

Universidade Federal do Paraná

Programa de Pós-Graduação em Métodos Numéricos em Engenharia

marisabaptis@yahoo.com.br

Luiz Homero Bastos Cunico

Universidade Federal do Paraná

Programa de Pós-Graduação em Métodos Numéricos em Engenharia

Maria Teresinha Arns Steiner

Universidade Federal do Paraná

Departamento de Matemática / Programa de Pós-Graduação em Métodos
Numéricos em Engenharia

Resumo: O presente trabalho tem por objetivo a utilização da técnica de Redes Neurais Artificiais na estimação dos valores venais de imóveis urbanos da cidade de Guarapuava, PR. Para tanto, utilizaram-se dados do Cadastro Imobiliário fornecidos pelo setor de Planejamento da Prefeitura Municipal. O modelo inicial foi composto por treze variáveis/atributos do cadastro: bairro, setor, pavimentação, esgoto, iluminação pública, área do terreno, pedologia, topografia, situação, área edificada, tipo, estrutura e conservação. A técnica da Análise das Componentes Principais foi usada para reduzir e transformar as variáveis originais em nove fatores. As Redes Neurais Artificiais desenvolvidas foram do tipo *feedforward*, utilizando o algoritmo de treinamento *Levenberg-Marquardt*, com uma camada oculta. Os resultados da amostra de dados completa foram comparados com os resultados obtidos dividindo-se a mesma em grupos menores, compostos de bairros com características

semelhantes, sendo que os resultados obtidos com estes últimos (grupos menores) foram superiores aos obtidos com o primeiro (amostra completa).

Palavras-chave: redes neurais artificiais; análise multivariada; engenharia de avaliações.

Abstract: This paper considers the use of the Artificial Neural Networks technique in the market value estimation for urban real estate in the city of Guarapuava, PR. Official data from the real estate register were used, having been provided by the Planning Sector at the City Hall. The initial model encompassed thirteen variables/attributes from the registration: neighborhood, sector, pavement, sewage, public lighting, land area, soil, topography, status, built area, type, structure and maintenance condition. The Principal Component Analysis technique was used in order to reduce and to transform the original variables into nine factors. The Artificial Neural Networks developed were feedforward type, using the Levenberg Marquardt training algorithm with one hidden layer. The results from the complete data sample were compared with the results obtained by dividing the sample into small groups, consisting in neighborhoods with similar features; the results obtained from the latter (smaller samples), were better than the former (complete data sample).

Key words: artificial neural networks; multivariate analysis; engineering of evaluations.

1 Introdução

A Engenharia de Avaliações, cada vez mais, vem evoluindo no Brasil, através da ação de inúmeros profissionais e pesquisadores que se dedicam ao estudo e pesquisa de problemas da área, bem como à divulgação de técnicas para avaliação de imóveis.

Segundo Dantas (2003), a Engenharia de Avaliações é uma especialidade da engenharia que reúne um conjunto amplo de conhecimentos da área de engenharia e arquitetura, bem como de outras áreas das ciências sociais, exatas

e da natureza, com o objetivo de determinar tecnicamente o valor de um bem, de seus direitos, frutos e custos de produção. Mendonça et al. (1998) citam as situações mais frequentes em que a avaliação técnica se faz necessária, dividindo-as em três âmbitos. No âmbito particular, os autores citam o uso da avaliação de imóveis na compra e venda dos mesmos em que os interessados desejam ter uma ideia precisa do bem em questão; na reavaliação de ativos de empresas para efeito de aumento de capital; por empresas para atendimento ao contido na Lei das Sociedades Anônimas; em divisões, heranças e meações, e nas discussões acerca de lançamentos de impostos. No âmbito público, é citado o uso da avaliação de imóveis para fins de compra e privatização e no cálculo de valores para lançamentos de impostos. Finalmente, no âmbito judicial, é utilizada nas discussões entre pessoas físicas ou jurídicas que envolvam valores de imóveis, frequentes em ações demarcatórias, possessórias e indenizatórias e, também, nas discussões acerca de indenizações por desapropriações ou servidões de passagem.

Os métodos estatísticos são os procedimentos de modelagem matemática mais utilizados, atualmente, para a avaliação imobiliária. O modelo de Regressão Linear Múltipla é o preferido dos avaliadores, por ter se mostrado bastante eficiente, embora fatores tais como a complexidade dos modelos, dificuldades de implementação, excesso de variáveis/atributos envolvidos e desconhecimento da relação entre os mesmos, possam comprometer a análise dos resultados. A escolha de Redes Neurais Artificiais, para a estimação do valor de um imóvel, apresenta como vantagem a dispensa de linearidade entre as variáveis, se for o caso, e a facilidade de implementação.

2 Descrição do Problema e Obtenção dos Dados

Os dados referentes a esta pesquisa foram obtidos junto ao setor de planejamento da Prefeitura Municipal de Guarapuava, PR. Para o levantamento dos dados fez-se uma filtragem, separando para a análise trezentos imóveis residenciais (casas e apartamentos), com valores venais entre R\$ 40.000,00 a R\$ 200.000,00, localizados na área urbana do município mencionado. As treze variáveis escolhidas para cada imóvel foram, basicamente, as constantes no Boletim de Informações

Cadastrais, sendo elas: localização (bairro), setor, pavimentação, iluminação pública, esgoto, área do terreno, pedologia, topografia, situação, área edificada, tipo, estrutura, conservação e, ainda, o seu valor venal. As variáveis água e rede elétrica foram desconsideradas do modelo, pois todos os imóveis da amostra em questão estavam contemplados com estes itens. Além disso, a amostra foi reduzida de trezentos para 256 unidades, pois foram eliminados os imóveis repetidos como, por exemplo, um prédio com vários apartamentos com as mesmas características; neste caso, foi considerado apenas um apartamento.

3 Redes Neurais Artificiais

As Redes Neurais Artificiais (RNA's), amplamente utilizadas pela comunidade científica, têm o seu funcionamento inspirado na estrutura biológica do cérebro humano e podem ser entendidas como conjuntos bem estruturados de unidades de processamento, interligados por canais de comunicação, cada qual tendo um determinado peso correspondente a um valor numérico. No decorrer do processo de "aprendizagem", os referidos pesos vão se ajustando de forma a atingir o objetivo pretendido que, no presente trabalho, é a previsão dos valores venais dos imóveis.

3.1 Algoritmo de *Levenberg-Marquardt*

Em uma RNA de Múltiplas Camadas, o algoritmo de retro-propagação (*back-propagation*) padrão utiliza o método do gradiente decrescente no decorrer do processo de aprendizagem de forma a minimizar a função erro. Já o algoritmo de *Levenberg-Marquardt* (*LM*) utiliza uma aproximação pelo método de Newton; esta aproximação é obtida a partir do método de Gauss-Newton modificado. A atualização dos pesos através do método de Gauss-Newton é feita da seguinte forma (HAYKIN, 2001):

$$w(n+1) = w(n) - (J^T(n) J(n) + \delta I)^{-1} J^T(n) e(n)$$

onde n é o número da iteração; $w(n)$ é o peso na iteração n ; $J(n)$ é a matriz jacobiana do vetor do erro em relação ao vetor peso na iteração n ; δ é uma constante positiva pequena escolhida para assegurar que $(J^T(n)J(n) + \delta I)$ seja definida positiva para todo n e, finalmente, $e(n)$ é o erro na iteração n .

4 Análise das Componentes Principais

A Análise das Componentes Principais é uma das técnicas da Análise Estatística Multivariada utilizada para investigar a estrutura de covariância em um conjunto de p variáveis correlacionadas. Para tanto, transforma-se o conjunto das p variáveis originais em um novo conjunto com cardinal

$k < p$ de variáveis não-correlacionadas chamadas de Componentes Principais, tendo propriedades especiais em termos de variâncias. Assim, a Análise das Componentes Principais explica a estrutura de variância e covariância de um vetor aleatório através de poucas combinações lineares das variáveis originais. O objetivo geral consiste, tanto em reduzir os dados como em facilitar a interpretação, pois trata-se de uma transformação ortogonal de eixos, tornando as novas variáveis (combinações lineares) não correlacionadas (JOHNSON e WICHERN, 2002) e nas direções com maior variabilidade.

A reprodução da variabilidade total do sistema requer as p variáveis, porém, frequentemente a maior parte dessa variabilidade pode ser explicada por um número pequeno $k < p$, de componentes principais. Neste caso, existe praticamente a mesma quantidade de informações nas k componentes principais do que nas p variáveis originais.

4.1 Componentes Principais da Amostra

Geralmente a estrutura de covariância, resumida na matriz de covariância, Σ , é desconhecida, então a obtenção das componentes principais é feita a partir de estimativas da referida matriz, S , ou da matriz de correlação, R . Então, obtém-se as estimativas dos elementos da estrutura de covariância do vetor aleatório \underline{X} , ou seja, os autovalores e os correspondentes autovetores e constroem-se as componentes

principais amostrais. As propriedades das componentes principais se mantêm e são obtidas com base em estimadores.

A obtenção das componentes principais com base nas informações da matriz de correlação é preferida, pelo fato de se conseguir eliminar o efeito de escala nos valores das componentes do vetor de variáveis originais X . Vale salientar que a matriz de correlação é uma matriz de covariância, mas de variáveis padronizadas, assim consegue-se eliminar a influência da escala na magnitude das variáveis (JOHNSON e WICHERN, 2002).

Os autovetores definem as direções da máxima variabilidade e os autovalores especificam as variâncias. Quando alguns autovalores são muito maiores que os demais, significa que a maior parte da variância total pode ser explicada por um número muito menor de componentes principais, do que as p dimensões do vetor X .

5 Implementação das Técnicas para Resolução do Problema

Após o levantamento dos dados, implementaram as técnicas propostas neste trabalho. Para a avaliação estatística das características dos dados, foi realizada a Análise das Componentes Principais, que teve como propósito, avaliar a importância relativa das variáveis que compõem a amostra de dados. Este procedimento pode ser adotado em situações onde o número de variáveis (entradas) do modelo é grande e, além disso, as mesmas são altamente correlacionadas. Esta análise possibilita a redução da dimensão dos vetores de entrada. Para a realização da Análise das Componentes Principais, foi utilizado o *software Estatística 5.0*.

Aplicando-se a técnica da Análise das Componentes Principais no conjunto de dados originais composto de treze variáveis (bairro, setor, pavimentação, esgoto, iluminação, área do terreno, pedologia, topografia, situação, área edificada, tipo, estrutura e conservação), percebeu-se que um modelo com apenas nove componentes principais seria suficiente para representar a estrutura de covariância inicial, com 18,69% de perda de informações. Aceitaram autovalores maiores que 0,75. Porém, utilizando o critério de Kaiser (JOHNSON e WICHERN, 2002), o número de variáveis reduziria para seis, entretanto a perda de informações seria

relativamente grande, em torno de 38%. Na tabela1 observa-se a proporção da variância explicada pelos autovalores ordenados de forma decrescente.

Tabela 1. Autovalores em ordem decrescente de magnitude

	Autovalores	% Total Variância	Cumul. Autovalores	Cumul. %
1	1,858557	14,29659	1,85856	14,29659
2	1,511854	11,62965	3,37041	25,92624
3	1,296738	9,97491	4,66715	35,90115
4	1,208301	9,29463	5,87545	45,19577
5	4,400844	8,46803	6,97629	53,66380
6	1,044903	8,03772	8,02120	61,70152
7	0,896430	6,89561	8,91763	68,59713
8	0,831275	6,39442	9,74890	74,99155
9	0,821521	6,31939	10,57042	81,31094
10	0,727890	5,59916	11,29831	86,91010
11	0,653833	5,02948	11,95215	91,93960
12	0,555787	4,27529	12,50793	96,21490
13	0,492067	3,78513	13,00000	100,0000

Fonte: Autores

A Análise das Componentes Principais pode ser usada para julgar a importância das próprias variáveis originais escolhidas. Assim, as variáveis originais com maior peso na combinação linear dos primeiros componentes principais são as mais importantes do ponto de vista estatístico. Como o objetivo é reduzir o número de variáveis (de treze para nove neste trabalho), preservando o máximo possível de informação sobre as entradas, esta técnica oferece um procedimento de aprendizagem auto-organizada.

5.1 Primeiro Experimento – Amostra Completa

Para a realização deste trabalho, foi escolhido o *software Matlab 6.5*. Este *software* oferece pacotes prontos para utilização das RNA's em seu *toolbox*.

A topologia empregada foi a Perceptron de Múltiplas Camadas (*MultiLayer Perceptron - MLP*) *feedforward* com uma camada de entrada, consistindo de nove nós, conforme as variáveis obtidas na Análise das Componentes Principais, uma camada escondida, e uma camada de saída com um único neurônio que fornecerá o valor venal do imóvel. O número de neurônios na camada oculta variou de nenhum neurônio até doze neurônios.

A função de ativação deve ser necessariamente contínua e diferenciável. Neste estudo, foi empregada, na camada oculta e na camada de saída, a função de ativação não linear do tipo sigmoideal (logsig), que assume um intervalo contínuo de valores entre 0 e 1 e é definida como uma função estritamente crescente, que exhibe um balanceamento adequado entre comportamento linear e não linear.

No treinamento foi utilizado o algoritmo de *LM*, para o ajuste dos pesos sinápticos do *Perceptron* de Múltiplas Camadas. De acordo com HAYKIN (2001), um *Perceptron* de Múltiplas Camadas treinado com o algoritmo de retro-propagação pode ser visto como um veículo prático para realizar um mapeamento não linear de entrada saída de natureza geral. Para a avaliação do desempenho da rede neural, utilizou-se o erro quadrático médio (*MSE – Mean Squared Error*), que é a média do somatório dos quadrados dos erros de cada caso, tanto do conjunto de treinamento, quanto do conjunto de testes. É usual, também, se adotar a raiz quadrada do erro quadrático médio (*rmse*) no cálculo do erro da rede.

Na continuidade, a amostra foi dividida em dois grupos. O primeiro grupo chamado de Conjunto de Treinamento (P) com 66% (170 imóveis) da amostra e o segundo com 34% (86 imóveis) foram separados em um Conjunto de Teste (S), tomando-se o cuidado de manter a mesma proporção de valores dos imóveis nos dois conjuntos. Foi escolhida uma nomenclatura para as redes de forma a representar, pela ordem, as seguintes características: número de parâmetros de entrada, número de neurônios na camada escondida e número de iterações do treinamento. Dessa forma, a denominação das redes deste trabalho começa com um número correspondente à quantidade de variáveis de entrada, seguido da letra E, seguida do número de neurônios na camada oculta e a letra N e número de neurônios na camada de saída seguido também pela letra N, e finalmente o número de iterações no treinamento. Por exemplo, a rede 9E8N1N50 é uma rede com nove

entradas, oito neurônios na camada oculta, um neurônio na camada de saída e foi treinada com cinquenta iterações. Cada conjunto foi treinado aproximadamente cinquenta vezes, variando-se os valores dos pesos iniciais, cujo critério de parada utilizado foi o número de iterações (épocas) para cada treinamento.

As simulações que apresentaram melhores resultados são mostrados no quadro 1 a seguir.

Quadro 1. Resultado das melhores simulações

Rede	Erro Médio Quadrático mse (%)	
	Treinamento	Teste
9E1N50	14,89	16,41
9E1N1N50	12,32	14,82
9E2N1N50	11,39	15,39
9E3N1N50	7,61	13,15
9E4N1N50	6,96	12,50
9E5N1N50	6,28	10,40
9E6N1N50	5,62	12,09
9E7N1N50	3,91	13,94
9E8N1N50	4,96	11,13
9E9N1N50	3,91	13,81
9E10N1N50	4,64	12,88
9E11N1N50	3,14	12,23
9E12N1N50	2,81	15,81

Fonte: Autores

O quadro 1 apresenta as redes consideradas de melhor desempenho, dentre as várias testadas, por apresentarem os menores percentuais de erros, tanto no treinamento como no teste. As redes apresentaram no treinamento valores para os *rmse's* no intervalo de 2,81% a 14,89% e os *rmse's* de teste ficaram no intervalo entre 10,40% e 16,41%, sendo que a rede com cinco neurônios na camada oculta apresentou o melhor resultado de teste, ou seja, melhor capacidade de generalização.

5.2 Segundo Experimento – Setor 9

Neste segundo experimento, inicialmente utilizaram-se os dados originais (13 variáveis). O primeiro grupo de dados, composto por 52 amostras foi denominado de Setor 9, que contém imóveis localizados em três bairros da cidade. O número de variáveis deste conjunto passou de treze para dez, uma vez que todos os imóveis são do Setor 9, todos os imóveis localizados nesta área possuem pavimentação e iluminação, podendo-se, dessa forma, excluir estas três variáveis. Trabalhou-se, assim, com as dez variáveis do modelo inicial (bairro, esgoto, área do terreno, pedologia, topografia, situação, área edificada, tipo, estrutura e conservação). O conjunto de dados do Setor 9 foi então dividido em dois conjuntos, sendo 75% da amostra (39 observações para o Conjunto de Treinamento da rede neural) e, 25% da amostra (treze observações para o Conjunto de Teste). Os resultados estão apresentados no quadro 2.

Quadro 2. Resultado das melhores simulações – Setor 9

Rede	Erro Médio Quadrático mse (%)	
	Treinamento	Teste
10E2N1N50	1,85	6,79
10E4N1N50	1,98	6,74
10E6N1N50	1,82	7,94
10E8N1N50	1,79	5,37
10E10N1N50	1,92	8,21
10E12N1N50	2,00	9,21

Fonte: Autores

O quadro 2 apresenta as redes consideradas de melhor desempenho por mostrarem os menores percentuais de erros, tanto no treinamento como no teste. As redes apresentaram no treinamento valores para os *rmse's* no intervalo de 1,79% a 2% e no teste os *rmse's* ficaram no intervalo entre 5,37% e 9,21%, sendo que a rede com oito neurônios na camada oculta apresentou o melhor resultado. Em seguida, realizou-se a Análise das Componentes Principais apenas para os imóveis pertencentes ao Setor 9, com o intuito de reduzir o número de variáveis. A tabela 2 mostra o resultado da análise.

Tabela 2. Autovalores em ordem decrescente de magnitude – Setor 9

	Autovalores	% Total Variância	Cumul. Autovalores	Cumul. %
1	1,842741	18,42741	1,842741	18,42741
2	1,705270	17,05270	3,548011	35,48011
3	1,541218	15,41218	5,089229	50,89229
4	1,129855	11,29855	6,219084	62,19084
5	1,022285	10,22285	7,241370	72,41370
6	0,859281	8,59281	8,100650	81,00650
7	0,607593	6,07593	8,708240	87,08240
8	0,541639	5,41639	9,249880	92,49880
9	0,441756	4,41756	9,691640	96,91640
10	0,308361	3,08361	10,000000	100,00000

Fonte: Autores

A Análise mostrou que seis variáveis seriam suficientes para representar a estrutura de covariância inicial, com 19% de perda de informações. O critério de corte utilizado foram autovalores acima de 0,75.

Assim sendo, foram treinadas redes com seis variáveis na camada de entrada. Os melhores resultados das simulações estão demonstrados no quadro 5.3.

Quadro 3. Resultado das melhores simulações, com análise multivariada nos dados – Setor 9

Rede	Erro Médio Quadrático mse (%)	
	Treinamento	Teste
6E2N1N50	0,46	7,25
6E4N1N50	0,31	5,34
6E6N1N50	0,29	5,26
6E8N1N50	0,29	5,95
6E10N1N50	0,31	8,17
6E12N1N50	0,28	7,30

Fonte: Autores

O quadro 3 apresenta as redes consideradas com melhor desempenho por mostrarem os menores percentuais de erros, tanto de treinamento como de teste. As redes apresentaram no treinamento *rmse's* no intervalo de 0,28% a 0,46% e os

rmse's de teste ficaram no intervalo entre 5,26% a 7,30%, sendo que a rede com seis neurônios na camada oculta apresentou o melhor resultado.

5.3 Terceiro Experimento – Setor 10

O terceiro experimento foi realizado sobre um segundo grupo de dados, denominado de Setor 10, que contém imóveis localizados em quatro bairros da cidade, composto por 79 amostras. Nesse grupo de dados, também aplicou-se a técnica de Redes Neurais Artificiais antes da aplicação da Análise das Componentes Principais. O número de variáveis desse conjunto passou de treze para oito, uma vez que todos os imóveis são do Setor 10, todos possuem pavimentação, esgoto e iluminação e a variável pedologia é a mesma para todos. Trabalharam as redes com as oito variáveis do modelo inicial (bairro, área do terreno, topografia, situação, área edificada, tipo, estrutura e conservação). O conjunto de dados do Setor 10 foi dividido em dois conjuntos, sendo 75% da amostra, ou seja, 59 observações para o Conjunto de Treinamento da rede e 25% da amostra, ou seja, vinte observações para o Conjunto de Teste. Os resultados estão apresentados no quadro 4.

Quadro 4. Resultado das melhores simulações – Setor 10

Rede	Erro Médio Quadrático mse (%)	
	Treinamento	Teste
8E2N1N50	6,86	16,22
8E4N1N50	5,98	17,24
8E6N1N50	3,31	11,11
8E8N1N50	4,62	22,25
8E10N1N50	1,83	21,84
8E12N1N50	3,21	22,46

Fonte: Autores

O quadro 4 apresenta as redes consideradas com melhor desempenho por mostrarem os menores percentuais de erros, tanto de treinamento como de teste. As redes apresentaram no treinamento valores de *rmse's* no intervalo de 1,83% a 6,89% e no teste os valores de *rmse's* ficaram no intervalo entre 11,11% e 22,46%,

sendo que a rede com seis neurônios na camada oculta apresentou o melhor resultado. Realizou-se na sequência, a Análise das Componentes Principais, com o intuito de reduzir o número de variáveis. A tabela 3 mostra o resultado dessa análise, quanto à escolha dos fatores.

Tabela 3. Autovalores em ordem decrescente de magnitude – Setor 10

	Autovalores	% Total Variância	Cumul. Autovalores	Cumul. %
1	2,184209	27,30261	2,184209	27,3026
2	1,436104	17,95130	3,620313	45,2539
3	1,092079	13,65098	4,712392	58,9049
4	0,948262	11,85328	5,660654	70,7582
5	0,824562	10,30703	6,485216	81,0652
6	0,723396	9,04245	7,208612	90,1076
7	0,495078	6,18848	7,703690	96,2961
8	0,296310	3,70387	8,000000	100,0000

Fonte: Autores

A Análise mostrou que seis variáveis seriam suficientes para representar a estrutura de covariância inicial, com 9,89% de perda de informações. O critério de corte utilizado, foi autovalores acima de 0,70. Realizou-se o treinamento da rede após a Análise das Componentes Principais. Os melhores resultados das simulações estão apresentados no quadro 5, com somente seis variáveis na camada de entrada e número de neurônios variados na camada oculta.

Quadro 5. Resultado das melhores simulações, com análise multivariada dos dados – Setor 10

Rede	Erro Médio Quadrático	
	Mrse (%)	
	Treinamento	Teste
6E2N1N50	11,74	13,90
6E4N1N50	4,30	18,58
6E6N1N50	1,88	18,85
6E8N1N50	1,78	20,30
6E10N1N50	0,43	22,24
6E12N1N50	0,31	22,96

Fonte: Autores

O quadro 5 apresenta as redes consideradas com melhor desempenho por mostrarem os menores percentuais de erros, tanto de treinamento como de teste. As redes apresentaram no treinamento valores do *rmse's* no intervalo de 0,31% a 11,74% e no teste os valores do *rmse's* ficaram no intervalo entre 13,90% e 22,96%, sendo que a rede com seis neurônios na camada oculta apresentou o melhor resultado.

6. Conclusões

Este trabalho teve como objetivo estudar a técnica de RNA's para determinação de valor venal de imóveis na cidade de Guarapuava, PR. Utilizando os dados do Cadastro Imobiliário Municipal, aplicou-se a referida técnica, fazendo uso do algoritmo matemático de *Levenberg Marquardt* com uma camada oculta.

O algoritmo utilizado exige pouco tempo computacional, necessitando de um número relativamente pequeno de iterações (no trabalho utilizaram 50 iterações em cada um dos treinos realizados), além do *software (Matlab)* que faz uso de Redes Neurais, apresentar uma forma bastante acessível de interface com o usuário, facilitando a aplicação.

A Análise Multivariada, através da Análise das Componentes Principais, reduz o número de variáveis, não prejudicando a compreensibilidade do modelo obtendo variáveis não correlacionadas. Neste estudo somente foram utilizadas as variáveis constantes no Cadastro Técnico Imobiliário do Município (treze variáveis), que são relativamente poucas, para se avaliar corretamente um imóvel mas, mesmo assim, a aplicação da técnica da Análise das Componentes Principais, mostrou que é possível reduzir ainda mais este número mantendo as informações.

Os resultados obtidos através das RNA's, quando da separação por setores, foram melhores em relação ao modelo contendo todos os dados, ou seja, quando setorizou-se a amostra, tanto os erros de treinamento quanto de testes, foram significativamente melhores. É possível observar, também, que a Análise das Componentes Principais melhorou os resultados de treinamento e de testes para a análise do Setor 9. Já no Setor 10, o erro de teste foi superior após a realização da Análise das Componentes Principais.

Além disso, é importante observar que outras variáveis adicionais como, acessibilidade, distância do centro comercial, acesso a escolas, hospitais, supermercados também são importantes tanto na determinação da aquisição de um imóvel quanto para a sua valorização e, desta forma, estas variáveis mesmas poderiam também ser levadas em consideração no modelo também melhorando, ainda mais, a previsão dos valores venais de imóveis.

7. Referências

DANTAS, R. A. *Engenharia de Avaliações: uma introdução à metodologia científica*. São Paulo: Pini, 2003.

HAYKIN, S. *Redes Neurais: princípios e prática*. 2. ed. Porto Alegre: Bookman, 2001.

JOHNSON, R. A.; WICHERN, D.W. *Applied Multivariate Statistical Analysis*. 5. ed. Nova Jersey: Prentice Hall, Inc., 2002

MENDONÇA, M.C.; SOLLERO FILHO, M.; CURI, E; AGUIAR, J.B.; QUEIROGA, H.S.; MAIA, E. A.; AQUINO, R.; RESENDE, O.; CANÇADO, J.M.M. *Fundamentos de avaliações patrimoniais e perícias de engenharia*. São Paulo: Pini, 1998.