

## **Determinação do intervalo de confiança em redes neurais artificiais**

**Benvenho, Agnaldo Calvi<sup>1</sup>**

<sup>1</sup> Autônomo, Rua Maestro Fco. Fortunato, 786, cjto. 84, Presidente Prudente, SP, Brasil,  
[abenvenho@terra.com.br](mailto:abenvenho@terra.com.br)

### **RESUMO**

As redes neurais artificiais são uma metodologia emergente na avaliação de bens, em geral, e de imóveis, em particular. No entanto, o emprego desta técnica fornece estimativas pontuais de valor. É importante, na maioria das aplicações, especialmente na avaliação, uma estimativa da confiabilidade do valor calculado, o intervalo de confiança. Neste artigo, será apresentada uma metodologia para o cálculo do intervalo de confiança de um valor pontual estimado, considerando a rede neural inferida como uma regressão não linear múltipla.

**Palavras-chave:** intervalo de confiança, redes neurais artificiais.

## Confidence interval determination on artificial neural networks

### ABSTRACT

Artificial neural networks are an emergent methodology on asset appraisal, in general, and real estate, in particular. However, this technique provides a punctual estimative of value. It is important, in most applications, specially in appraisal, a reliability estimate of the calculated value, the confidence interval. In this paper, a methodology to calculate confidence interval for an estimated punctual value will be presented, considering the inferred neural network as a multiple nonlinear regression.

**Key-words:** confidence intervals, artificial neural networks.

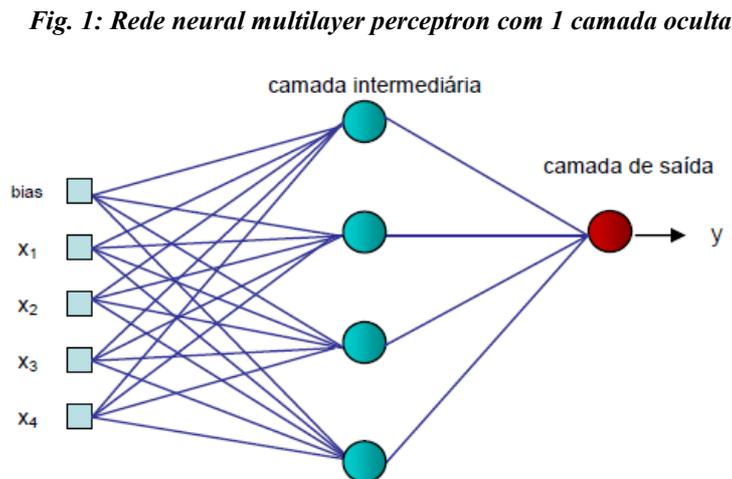
## 1. REDES NEURAIS ARTIFICIAIS

Segundo Gonzáles e Formoso (2000), redes neurais artificiais são modelos computacionais que buscam emular o cérebro humano em sua forma de processar informações. Elas são estruturalmente compostas por um conjunto de múltiplos processadores (unidades), cada qual tendo uma pequena quantidade de memória, ligados entre si por canais de comunicação (conectores), que tem a capacidade de transmitir dados numéricos, codificados em vários formatos (sinais), à semelhança do neurônio humano. As unidades operam apenas seus dados locais e suas entradas ocorrem pelas conexões.

O modelo básico de uma rede neural artificial trata-se de uma série de neurônios de entrada (camada de entrada), os quais são responsáveis pelo primeiro processamento, utilizando funções matemáticas e pesos e por um ou mais neurônios de saída (camada de saída), onde são passadas as variáveis resposta.

O modelo de rede neural artificial mais difundido é o *perceptron* de múltiplas camadas, também conhecido como *multilayer perceptron*. Este modelo consiste na utilização de camadas de neurônios artificiais, sendo uma de entrada, uma de saída e uma série de intermediárias, também conhecidas como *hidden layers*. Dependendo do tipo de problema, pode-se usar diversas camadas intermediárias, muito embora, de acordo com o Teorema de Kolmogorv-Nielsen, qualquer função não linear contínua pode ser aproximada por uma rede *multilayer perceptron* com uma camada intermediária.

A Fig. 1 mostra uma configuração básica de rede neural *multilayer perceptron* com uma camada intermediária



Para as ligações entre dois neurônios  $i$  e  $j$ , teremos um peso específico associado a conexão,  $w_{ij}$ . Para cada neurônio, existe uma função de ativação. No caso das *multilayer perceptron*, as mais usadas são a linear,  $y = x$  e a sigmóide logística, representada pela Eq. 1.

$$y = \frac{1}{1 + e^{-x}} \quad (1)$$

Segundo Guedes (2001), é necessário treinar a rede neural, ou seja, apresentam-se pares de dados de entrada e saída, permitindo a ela fazer associações descobrindo assim a existência de algum padrão de comportamento.

O treinamento de uma rede *multilayer perceptron* é feito pelo algoritmo de retropropagação, também conhecido com regra delta generalizada. Ela consiste em ajustar os pesos  $w_{ij}$ , conforme as entradas  $x_p$  ( $x_1, x_2, \dots, x_p$ ) e saídas  $y_k$  ( $y_1, y_2, \dots, y_k$ ). Uma descrição detalhada pode ser encontrada em Guedes (2001), Azevedo, Brasil e Oliveira (2000) e Braga, Carvalho e Ludemir (2003).

## 2. INTERVALO DE CONFIANÇA

### 2.1 Conceito

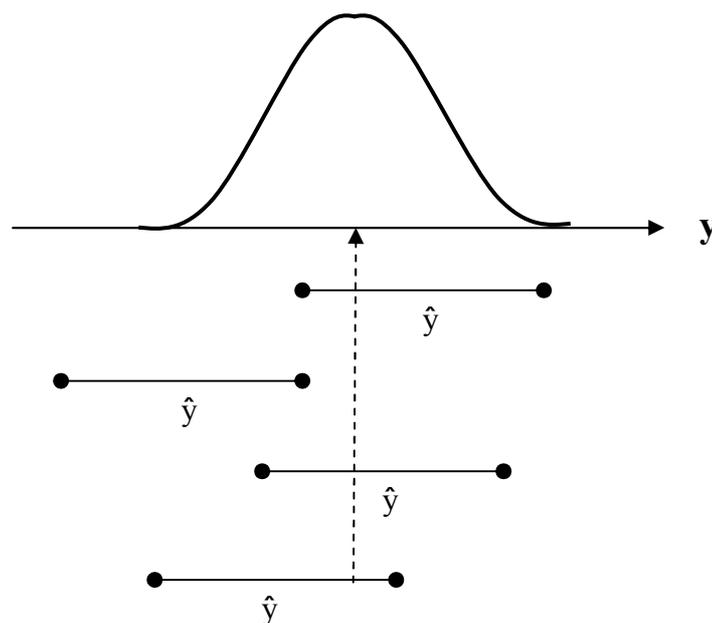
O intervalo de confiança é um intervalo estimado para um parâmetro estatístico (por exemplo, a média) e trata-se de um indicador de confiabilidade de uma estimativa. Em sentido estrito, o intervalo de confiança para um parâmetro populacional é um intervalo com proporção  $p$  associada (sendo  $0 \leq p \leq 1$ ), gerada por uma amostra aleatória de uma população subjacente, de tal forma que se a amostra for repetida inúmeras vezes e o intervalo de confiança for recalculado para cada amostra através do mesmo método, uma proporção  $p$  dos intervalos de confiança conterá o parâmetro estatístico em análise.

De acordo com Pires (2000), para uma amostra com média e variância conhecidas, sendo  $\bar{x}$  a média observada de uma amostra aleatória de dimensão  $n$  de uma população normal (ou de uma população qualquer desde que  $n$  seja grande, sendo que neste caso o intervalo de confiança será aproximado), com variância conhecida  $\sigma^2$ , um intervalo de confiança a  $100 \times (1 - \alpha)\%$  para  $\mu$  (o verdadeiro valor do parâmetro estimado) é dado pela Eq. 2:

$$\bar{x} - a \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + a \frac{\sigma}{\sqrt{n}}, \text{ com } a : P(Z > a) = \frac{\alpha}{2}, \text{ sendo } a = z_{\alpha/2} \quad (2)$$

A Fig. 2, extraída de Dybowski e Roberts (2001) ilustra o conceito de intervalo de confiança.

**Fig. 2: Intervalo de confiança amostral em uma população**



## 2.2. Intervalo de Confiança para Regressão Linear

A regressão linear é uma técnica estatística que busca relacionar uma determinada variável – denominada variável dependente – com uma série de variáveis – denominadas independentes – que guardam com aquela uma relação de pertinência, de acordo com Eq. 3:

$$y_i = w_1x_{i1} + w_2x_{i2} + w_3x_{i3} + \dots + w_kx_{ik} + e_i, \quad i = 1, 2, \dots, n \quad (3)$$

sendo  $y_i$  a variável dependente  $B_j$  os regressores a serem determinados e  $w_{ik}$  as variáveis independentes ou explicativas.

A teoria e as técnicas de regressão linear são descritas em Dantas (1999) e Gujarati (2000).

Considerando-se um valor estimado  $\hat{y}_h = w_1x_{1h} + w_2x_{2h} + \dots + w_kx_{kh}$ , sendo  $\hat{y}_h$  o valor estimado para  $y_h$ , o vetor  $(x_{1h}, \dots, x_{kh})$  dos valores para cada uma das variáveis independentes e considerando  $s(\hat{y}_h)$  o desvio padrão relativo a estimativa de  $\hat{y}_h$ ,  $S_1$  o desvio padrão dos resíduos do modelo,  $n$  o número de amostras utilizado na determinação do modelo,  $s(w_i)$  o desvio padrão do regressor  $i$ ,  $\bar{x}_{hi}$  as médias para cada variável independente e  $cov(w_i, w_l)$ , sendo  $(l, i = 1, 2, \dots, k, l < i)$ , teremos as Eq. 4 e 5:

$$S(\hat{y}_h) = \sqrt{\frac{S_1^2}{n} + \sum_{j=i}^k (x_{hi} - \bar{x}_{hj})^2 s(w_j) + 2 \sum_{l < j}^k (x_{hl} - \bar{x}_{hl})(x_{hj} - \bar{x}_{hj}) cov(w_j, w_l)} \quad (4)$$

Em notação matricial, temos:

$$S(\hat{y}_h) = S_e \sqrt{\frac{1}{n} + (X - \bar{X})^T (X^T X)^{-1} (X_h - \bar{X})} \quad (5)$$

## 2.3. Intervalo de Confiança para Regressão Não Linear

Uma função de regressão do tipo  $y = f(x, \theta)$  quando a função que apresenta os melhores estimadores para  $y$ , tal que:  $y = f(x_i, \hat{\theta}) + e_i$ , sendo  $i = 1, 2, \dots, n$ ,  $e_i$  é um termo estocástico de erro e  $\hat{\theta}$  é não linear, ou seja, a função  $y$  não pode ser representada pela Eq. 6.

$$y = f(x_i, \hat{\theta}) = \sum_{j=1}^p x_{ij} \hat{\theta}_j, \quad \text{com } j = 1, 2, \dots, p. \quad (6)$$

De acordo com Donaldson e Schnabel (1985), existem três métodos para se determinar o intervalo de confiança para uma regressão não linear:

- Linearização dos regressores;
- Máxima verossimilhança;
- Bootstrap.

Ainda de acordo com os autores, o método da linearização dos regressores é o mais frequentemente utilizado, pois requer menor poder computacional e produz resultados de mais fácil interpretação. Todavia, produz subestimativas dos intervalos de confiança. Os outros

métodos, apesar de apresentar estimativas melhores, requerem grande poder de computação e produzem resultados que muitas vezes são difíceis de interpretar.

Neste âmbito, estudaremos o método da linearização para a determinação de intervalos de confiança m regressões lineares. Ainda, na linha dos autores, quando a determinação dos parâmetros  $\hat{\theta} \left( \hat{\theta} = (X^T X)^{-1} X^T Y \right)$ , sendo X a matriz das variáveis independentes, Y a matriz das variáveis dependentes e as notações T indicando matriz transposta e -1 matriz inversa, é feita pelo método dos mínimos quadrados não lineares, existem três variantes do método da linearização, um utilizando o operador Jacobiano, um utilizando o operador Hessiano e outro usando uma combinação dos dois. Donaldson e Schnabel (1985) concluem que o método que utiliza o operador Jacobiano é preferível, pois é mais simples, demanda menor poder computacional, é numericamente mais estável e tão preciso quanto os outros métodos.

Partindo desta conclusão, utilizaremos o primeiro método como forma de estimar o intervalo de confiança em uma regressão não linear.

De acordo com Chryssolouris, Lee e Ramsey (1996), considerando um modelo para um sistema qualquer  $f(x, \theta^*)$ , sendo x as entradas do modelo (variáveis independentes) e  $\theta^*$  representa os verdadeiros valores dos estimadores  $\theta$ . Assume-se que o erro e associado ao modelo é independente e, com variância  $\sigma^2$  e possui distribuição normal tal que  $N(0, \sigma^2)$ . Para n observações, tais que  $i = 1, 2, \dots, n$ , o sistema é representado pela Eq. 7:

$$y_i = f(x_i, \theta^*) + e_i, i = 1, 2, \dots, n \quad (7)$$

A estimativa por mínimos quadrados de  $\theta^*$  é  $\hat{\theta}$ , obtida através da minimização da função de erro, expressa na Eq. 8.

$$S(\theta) = \sum_{i=1}^n [y_i - f(x_i, \theta)]^2 \quad (8)$$

$$\text{onde } \hat{y}_i = f(x_i, \hat{\theta})$$

Sendo o modelo um bom estimador do sistema, então teremos  $\hat{\theta} \approx \theta^*$ . Logo, podemos linearizar o modelo aplicando a série de Taylor nas imediações de  $\theta^*$ , de acordo com as Eqs. 9 e 10.

$$f(x_i, \hat{\theta}) \approx f(x_i, \theta^*) + f_o^T (\hat{\theta} - \theta^*) \quad (9)$$

$$f_o^T = \left( \frac{\partial f(x_i, \theta^*)}{\partial \theta_1^*}, \frac{\partial f(x_i, \theta^*)}{\partial \theta_2^*}, \dots, \frac{\partial f(x_i, \theta^*)}{\partial \theta_n^*} \right) \quad (10)$$

Derivado das expressões acima, o erro pontual entre o valor do y observado e do  $\hat{y}$  estimado é dado pelas Eqs. 11 e 12.

$$y_o - \hat{y}_o \approx y_o - f(x_o, \theta^*) - f_o^T (\hat{\theta} - \theta^*) = e_o - f_o^T (\hat{\theta} - \theta^*) \quad (11)$$

$$E[y_o - \hat{y}_o] \approx E[e_o] - \mathbf{f}_o^T E[(\hat{\theta} - \theta^*)] \approx 0 \quad (12)$$

Considerando a independência estatística entre  $\hat{\theta}$  e o termo de erro, a Eq. 13 exprime variância esperada entre  $y$  e  $\hat{y}$ .

$$\text{var}[y_o - \hat{y}_o] = \text{var}[e_o] - \text{var}[\mathbf{f}_o^T (\hat{\theta} - \theta^*)] \quad (13)$$

A matriz Jacobiana  $F \cdot (\hat{\theta})$ , de acordo com Ungar, De Veaux e Rosengarten (1994), é representada pela Eq. 14

$$F \cdot (\hat{\theta}) = \left[ \frac{\partial f_i(x_i, \hat{\theta})}{\partial \hat{\theta}_j} \right] \quad (14)$$

ou sua expansão, de acordo com Chryssolouris, Lee e Ramsey (1996), acarreta na Eq. 15

$$F \cdot (\hat{\theta}) = \begin{bmatrix} \frac{\partial f_1(x_1, \hat{\theta})}{\partial \hat{\theta}_1} & \frac{\partial f_1(x_1, \hat{\theta})}{\partial \hat{\theta}_2} & \dots & \frac{\partial f_1(x_1, \hat{\theta})}{\partial \hat{\theta}_p} \\ \frac{\partial f_2(x_2, \hat{\theta})}{\partial \hat{\theta}_1} & \frac{\partial f_2(x_2, \hat{\theta})}{\partial \hat{\theta}_2} & \dots & \frac{\partial f_2(x_2, \hat{\theta})}{\partial \hat{\theta}_p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(x_n, \hat{\theta})}{\partial \hat{\theta}_1} & \frac{\partial f_n(x_n, \hat{\theta})}{\partial \hat{\theta}_2} & \dots & \frac{\partial f_n(x_n, \hat{\theta})}{\partial \hat{\theta}_p} \end{bmatrix} \quad (15)$$

Segundo Chryssolouris, Lee e Ramsey (1996), o erro  $e_o$  pode ser aproximado por uma distribuição normal tal que  $e_o \sim N(0, \sigma^2 I)$  onde  $I$  é a matriz identidade e o termo  $\hat{\theta}$  pode ser aproximado por uma distribuição normal tal que  $(\hat{\theta} - \theta^*) \sim N(0, \sigma^2 [F \cdot (\hat{\theta})^T F \cdot (\hat{\theta})]^{-1})$ , então  $\text{var}(y_o - \hat{y}_o) \approx \sigma^2 + \sigma^2 \mathbf{f}_o^T (F^T \cdot F)^{-1} \mathbf{f}_o$ . Também, segundo os mesmos, temos um estimador  $s^2$  para a variância  $\sigma^2$  que segue uma distribuição t de Student é calculado pela Eq. 16

$$s^2 = \frac{\|y - f(x, \hat{\theta})\|^2}{n - p} \quad (16)$$

sendo  $n$  é o número de amostras e  $p$  o número de parâmetros  $\hat{\theta}$ .

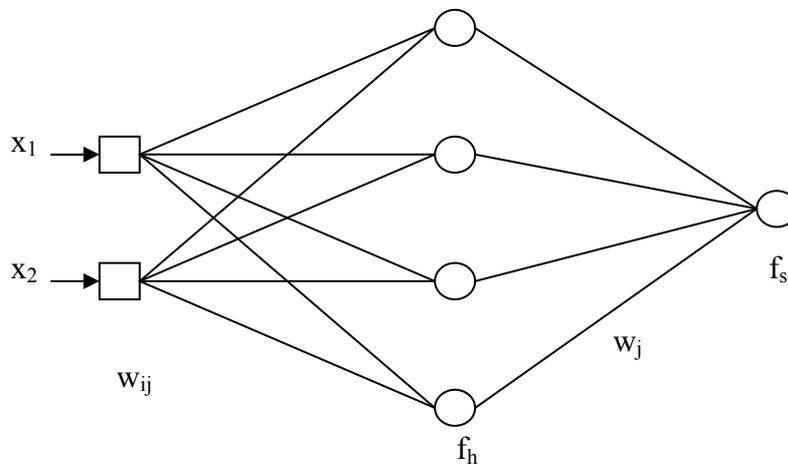
Assim, teremos um intervalo de confiança IC, considerando uma distribuição t de Student com  $n - p$  graus de liberdade e grau de confiança  $\alpha$  é calculado pela Eq. 17.

$$IC = t_{n-p}^{\alpha/2} s (1 + \mathbf{f}_0^T (F^T \cdot (\hat{\theta}) F \cdot (\hat{\theta}))^{-1} \mathbf{f}_0)^{\frac{1}{2}} \quad (17)$$

### 3. INTERVALO DE CONFIANÇA PARA REDES NEURAIS ARTIFICIAIS

Apresentamos uma rede neural artificial do tipo multilayer perceptron constituída por três camadas (entrada, intermediária e saída), com dois neurônios de entrada ( $x_1$ ,  $x_2$ ), quatro neurônios na camada intermediária ( $h_1$ ,  $h_2$  e  $h_3$  e  $h_4$ ) e um neurônio de saída ( $y$ ). Na Fig. 3, incluímos a representação dos pesos sinápticos ( $w_{ij}$ ) e as funções de ativação dos neurônios ( $f$ ).

*Figura 3: Rede neural multilayer perceptron com uma camada intermediária*



A rede neural esquematizada pode ser representada pela Eq. 18.

$$y = f_s \left( \sum_{j=1}^4 w_j f_h \left( \sum_{i=1}^2 w_{ij} x_i \right) \right) \quad (18)$$

Como podemos observar, a função  $y$  é dependente das entradas  $x_i$  e dos pesos,  $w_i$  e  $w_j$ . Estes, por sua vez, são não lineares, uma vez que variam conforme as equações  $f_h$  e  $f_s$ . Logo, os pesos  $w$  são equivalentes aos parâmetros  $\hat{\theta}$  da regressão não linear apresentada na seção anterior. Segundo Oliveira, Montini e Bergman (2007), métodos que fornecem intervalos de predição assintóticos para regressão não linear podem, em teoria, ser aplicados diretamente a redes neurais. Desta forma, podemos calcular o intervalo de confiança para uma estimativa feita pela função  $y = f(x, w)$  através da expressão apresentada previamente.

A rede pode apresentar várias camadas ocultas. Escolhemos uma rede com três camadas pela simplicidade do tratamento matemático, bem como pela consequência do Teorema de Kolmogorov-Nielsen, onde a mesma pode aproximar qualquer função não linear contínua e diferenciável, ao menos localmente, o que decorre que a rede tem capacidade de modelar o mercado (no nosso caso, imobiliário).

Considerando as que as funções de ativação sejam a linear e a sigmóide logística, dadas pelas Eqs. 19 e 20

$$f_s(x) = x \quad (19)$$

$$f_h(x) = \frac{1}{1 + e^{-x}} \quad (20)$$

obtemos a função resposta para Y, em função dos pesos e das entradas, acordo com Eq. 21:

$$Y = \frac{w_1}{1 + e^{-(w_{11}X_1 + w_{21}X_2)}} + \frac{w_2}{1 + e^{-(w_{12}X_1 + w_{22}X_2)}} + \frac{w_3}{1 + e^{-(w_{13}X_1 + w_{23}X_2)}} + \frac{w_4}{1 + e^{-(w_{14}X_1 + w_{24}X_2)}} \quad (21)$$

O modelo acima pode ser utilizado para calcular as saídas y dadas as variáveis de entrada,  $x_1$  e  $x_2$ . Busca-se então os pesos sinápticos  $w_1, w_2, w_3, w_4, w_{11}, w_{12}, w_{13}, w_{14}, w_{21}, w_{22}, w_{23}, w_{24}$  que forneçam a melhor aproximação para a variável de saída.

Consideremos o conjunto de dados apresentado na Tab. 1.

**Tabella 1: Dados para treinamento de rede neural**

y	$x_1$	$x_2$
473,68	2,00	665,00
600,00	3,00	600,00
675,00	3,00	600,00
675,00	3,00	600,00
670,53	3,00	604,00
670,53	3,00	604,00
691,75	3,00	618,00
750,00	3,00	630,00
261,90	2,00	1.890,00
951,34	3,00	596,00
983,19	3,00	595,00
898,62	3,00	651,00
504,40	3,00	1.249,00
1.270,59	4,00	595,00
85,96	1,00	8.900,00
1.554,28	4,00	608,00
1.813,95	4,00	645,00
3.450,00	5,00	600,00
225,00	1,00	10.000,00
288,00	1,00	10.000,00
1.598,89	3,00	2.533,00
1.928,57	4,00	2.100,00
2.571,43	4,00	2.100,00
424,53	2,00	742,00
680,67	3,00	595,00
639,20	3,00	704,00
1.111,76	3,00	425,00
780,99	3,00	605,00
832,46	3,00	573,00
831,93	3,00	595,00

Faz-se o treinamento, obtendo os pesos sinápticos apresentados na Tab. 2 a seguir.

**Tab. 2: Pesos sinápticos da rede**

Peso	Valor
w1	-0,355948
w2	0,238536
w3	0,441313
w4	-0,416180
w11	0,455222
w12	-0,266306
w13	-0,443213
w14	-0,350105
w21	0,424081
w22	0,583436
w23	0,168560
w24	-0,173391

Tomando como ponto de partida a Eq. 22, calculando-se as derivadas necessárias para as Eq. 10 e 16, obtemos a forma analítica da expressão do Intervalo de Confiança. Como trata-se de um intervalo para uma estimativa pontual, adotamos  $x_1 = 3$  e  $x_2 = 1.000,00$ , resultando em um  $y = 936,64$ .

Tendo-se os dados de treinamento é possível calcular  $s$  pela Eq. 17, sendo  $s = 232,42$ . Tomando estes, bem como os valores  $x_1$  e  $x_2$  da estimativa pontual, é possível calcular o termo  $f_0^T (F^T \cdot (\hat{\theta}) F \cdot (\hat{\theta}))^{-1} f_0$  e, por consequência, o termo  $(1 + f_0^T (F^T \cdot (\hat{\theta}) F \cdot (\hat{\theta}))^{-1} f_0)^{\frac{1}{2}}$ , cujo valor é 1,1547.

Considerando-se um nível de confiança de 95%, determina-se o termo  $t_{\frac{\alpha}{2}, n-p}$  pela distribuição t de Student, sendo o valor 0,461968.

Com os termos determinados, através da Eq. 18 chega-se ao intervalo de confiança para a estimativa pontual para as variáveis de entrada  $x_1 = 3$  e  $x_2 = 1.000,00$  e resultado  $y = 936,64$ .

$$IC = 123,98$$

$$y = 936,64 \pm 123,98$$

Na avaliação de um imóvel urbano, as variáveis  $x_1$  e  $x_2$  poderiam ser, por exemplo, a localização e área, sendo  $y$  o valor unitário. Em um imóvel rural, as mesmas poderiam ser o Índice Agrônômico e a área. Para uma máquina, poderíamos ter a potência e a idade.

#### 4. CONCLUSÕES

A determinação de intervalos de confiança em redes neurais artificiais é fundamental para analisar a confiabilidade de predições advindas das mesmas. Para o caso específico da engenharia de avaliações, as normas que regem os procedimentos avaliatórios tomam como base o intervalo de confiança para determinar a precisão da avaliação.

O método de cálculo exposto neste artigo demanda relativo esforço matemático, sobretudo no emprego de derivadas parciais de equações complexas e cálculo matricial. Também exige o conhecimento detalhado da arquitetura da rede treinada, inclusos aí número de camadas intermediárias, número de neurônios nas camadas, pesos sinápticos e funções de ativação.

No entanto, tais esforços são compensados, a medida que é possível inferir a confiabilidade da rede treinada com finalidade de avaliação, permitindo ao avaliador determinar se a mesma apresenta grau de precisão adequado.

O intervalo de confiança também se presta a analisar se a arquitetura da rede treinada, a qual é definida *a priori*, é adequada à resolução do problema e obtenção do valor perseguido.

Naturalmente, as hipóteses adotados no método apresentado podem trazer alguma distorção no cálculo dos intervalos de confiança, sobretudo na adoção da premissa de erros com distribuição normal e no cálculo de matrizes, onde pode haver necessidade de hipóteses adicionais. Entretanto, acreditamos que tais inconvenientes não o desmereçam, e, considerando que as redes neurais artificiais são uma metodologia emergente no campo da engenharia de avaliações, o cálculo do intervalo de confiança para valores preditos pelas mesmas é primordial para seu bom uso.

Enfim, existem dois problemas no cálculo do intervalo de confiança em redes neurais. O primeiro é inerente a formulação matemática das mesmas. Segundo Geman, Bienenstock e Doursat (1992), redes neurais com grande poder de generalização usualmente apresentam elevada variância, acarretando que, para boas estimativas pontuais, paga-se o preço de maior intervalo de confiança. Outro ponto, de acordo com Chryssolouris, Lee e Ramsey (1996), é de que ocorre uma subestimativa dos graus de liberdade em redes neurais, uma vez que as mesmas são treinadas com mais parâmetros do que o necessário.

Fica como recomendação para futuros trabalhos, o cálculo de intervalos de confiança em redes neurais utilizando os métodos bootstrap e de máxima verossimilhança e sua comparação com o método da regressão não linear.

## 5. BIBLIOGRAFIA

AZEVEDO, Fernando M.; Lourdes M. Brasil e Roberto C. L. Oliveira. Redes neurais com aplicação em controle e em sistemas especialistas. Florianópolis: Editora Bookstore; 2000;

BRAGA, Antônio; André C. de Carvalho e Teresa B. Ludemir. Redes neurais artificiais *in* REZENDE, Solange O. (organizadora). Sistemas inteligentes. Barueri: Editora Manole; 2003;

ARSLAN, Ozan e Oğuz Müftüoğlu. A practical method of confidence estimation for neural networks in image classification. Proceedings of International Conference on Intelligent Knowledge Systems (IKS-2004), 16-20 de Agosto, 2004;

CHRYSSOLOURIS, George; Moshin Lee e Alvin Ramsey. Confidence interval prediction for neural network models. IEEE Transactions on Neural Networks; Vol. 7, No. 1, Janeiro, 1996, pg. 229-232;

DANTAS, Rubens A. Engenharia de avaliações. Uma introdução a metodologia científica. São Paulo: Editora Pini; 1999;

DONALDSON, Janet R. e Robert Schnabel. Computational experience with confidence regions and confidence intervals for nonlinear least squares. Boulder: University of Colorado, Department of Computer Science; 1985;

DYBOWSKI, Richard e Stephen J. Roberts. Confidence intervals and prediction intervals for feed-forward neural networks *in* DYBOWSKI, R. e GANT, V. (organizadores). Clinical applications of artificial neural networks. Cambridge: Cambridge University Press; 2001;

GEMAN, Stuart; Elie Bienenstock e René Doursat. Neural networks and the bias/variance dilemma. *Neural Computation*, Vol. 4, 1992, pg. 1-58;

GONZÁLEZ, Marco A. S. e Carlos T. Formoso. Planta Genérica de Valores Utilizando Inteligência Artificial. In: CONGRESSO BRASILEIRO DE CADASTRO TÉCNICO MULTIFINALITÁRIO, 2000, Florianópolis. Anais eletrônicos...Florianópolis, 2000. Disponível em <[geodesia.ufsc.br/Geodesia-online/ARQUIVO](http://geodesia.ufsc.br/Geodesia-online/ARQUIVO)> Acesso em: 02 jun. 2003;

GUEDES, Jackson. Duas Ferramentas Poderosas a Disposição do Engenheiro de Avaliações – Modelos Lineares Generalizados e Redes Neurais. In XI CONGRESSO BRASILEIRO DE ENGENHARIA DE AVALIAÇÕES E PERÍCIAS, 2001, Guarapari. Anais... Vitória: IBAPE/ES, 2001. 1 CD;

GUJARATI, Damodar N. *Econometria Básica*. São Paulo: Editora Makron Books, 2000;

HWANG, Gene e Adam Ding. Prediction intervals for artificial neural networks. *Journal of the American Statistical Association*, Vol. 92, No. 438, Junho, 1997, pg. 748-757;

MAZUCHELI, Josmar e Jorge A. Achcar. Algumas considerações em regressão não linear. *Acta Scientiarum*. Vol. 24, No. 6, Novembro, 2002, pg. 1761-1770;

OLIVEIRA, Mauri A.; Alessandra de A. Montini e Daniel R. Bergman. Construção de intervalos de predição para redes neurais via regressão multivariada e sua aplicação em séries heterocedásticas. *Facef Pesquisa*, Vol. 10, No. 3, setembro, 2007, pg. 271-283;

PIRES, Ana. Notas de apoio às aulas de probabilidade e estatística, capítulo 7, IST, Outubro, 2000, disponível em <[www.math.ist.utl.pt/~apires/MaterialPE/AulaTCap7C.pdf](http://www.math.ist.utl.pt/~apires/MaterialPE/AulaTCap7C.pdf)> Acesso em: 15. jun. 2010;

RIVALS, Isabelle e Léon Personnaz. Construction of confidence intervals for neural networks based on least squares estimation. *Neural Networks*, Vol. 13, 2000, pg. 463-484;

UNGAR, Lyle H.; Richard D. de Veaux e Evelyn Rosengarten. Estimating prediction intervals for artificial neural networks. *Proceedings of the 9th Yale Workshop on Adaptive and Learning Systems*; 1996.