

R

# REDES NEURAIAS ARTIFICIAIS NA PREVISÃO DA INFLAÇÃO: APLICAÇÃO COMO FERRAMENTA DE APOIO À ANÁLISE DE DECISÕES FINANCEIRAS EM ORGANIZAÇÕES DE PEQUENO PORTE

## **LEONARDO AUGUSTO AMARAL TERRA**

*Mestre em Administração de Organizações pela Faculdade de Economia, Administração e*

*Contabilidade de Ribeirão Preto da Universidade de São Paulo (USP).*

*Professor do Curso de Administração da Faculdade de Educação São Luís.*

*Rua Floriano Peixoto, 839-873, Centro, Jaboticabal – SP – Brasil – CEP 14870-370*

*E-mail: leoterra@usp.br*

## **JOÃO LUIZ PASSADOR**

*Doutor em Administração pela Escola de Administração de Empresas de São Paulo da*

*Fundação Getúlio Vargas (FGV-SP).*

*Professor da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da*

*Universidade de São Paulo (USP).*

*Avenida dos Bandeirantes, 3.900, Monte Alegre, Ribeirão Preto – SP – Brasil – CEP 14040-905*

*E-mail: jlpassador@usp.br*

## RESUMO

As estimações das taxas de inflação são de fundamental importância para os gestores, pois as decisões de investimento estão intimamente ligadas a elas. Contudo, o comportamento inflacionário tende a ser não linear e até mesmo caótico, tornando difícil a sua correta estimação. Essa característica do fenômeno pode tornar imprecisos os modelos mais simples de previsão, acessíveis às pequenas organizações, uma vez que muitos deles necessitam de grandes manipulações de dados e/ou *softwares* especializados. O presente artigo tem por objetivo avaliar, por meio de análise formal estatística, a eficácia das redes neurais artificiais (RNA) na previsão da inflação, dentro da realidade de organizações de pequeno porte. As RNA são ferramentas adequadas para mensurar os fenômenos inflacionários, por se tratar de aproximações de funções polinomiais, capazes de lidar com fenômenos não lineares. Para esse processo, foram selecionados três modelos básicos de redes neurais artificiais *Multi Layer Perceptron*, passíveis de implementação a partir de planilhas eletrônicas de código aberto. Os três modelos foram testados a partir de um conjunto de variáveis independentes sugeridas por Bresser-Pereira e Nakano (1984), com defasagem de um, seis e doze meses. Para tal, foram utilizados testes de Wilcoxon, coeficiente de determinação  $R^2$  e o percentual de erro médio dos modelos. O conjunto de dados foi dividido em dois, sendo um grupo usado para treinamento das redes neurais artificiais, enquanto outro grupo era utilizado para verificar a capacidade de predição dos modelos e sua capacidade de generalização. Com isso, o trabalho concluiu que determinados modelos de redes neurais artificiais têm uma razoável capacidade de predição da inflação no curto prazo e se constituem em uma alternativa razoável para esse tipo de mensuração.

## PALAVRAS-CHAVE

Inflação; Redes neurais artificiais; *Perceptron*; Organizações de pequeno porte; Análise de decisão.

# 1 INTRODUÇÃO

São vários os métodos sugeridos para a análise de viabilidade e apoio às decisões financeiras. Contudo, Assaf Neto (2003, p. 33) observa que é importante que sejam levadas em consideração as questões inflacionárias relacionadas a esses modelos, pois a inflação promove discrepâncias no poder de compra da moeda com o passar do tempo, tornando necessária a exclusão de seus efeitos para viabilizar a mensuração do real poder de compra da moeda em determinado período. Apenas dessa forma os resultados esperados em um projeto poderiam ser comparados, livres da corrosão provocada pela perda de valor de compra da moeda (DAMODARAN, 1997, p. 227-229; VAN HORNE; WACHOWICZ JR., 1998, p. 29-30).

Além da própria problemática financeira, exposta anteriormente, Juster e Wachtel (1972) identificaram, por meio de seus estudos estatísticos, que o fantasma da inflação também atua como um custo de transação para os consumidores, que leva a perda de valor da moeda a ser um potencial fator inibidor do consumo. Com isso, a inflação tem um duplo impacto sobre as decisões de investimento, uma vez que afeta a mensuração do valor de projetos em determinada data e também a demanda esperada para ele. Assim, as estimações das taxas de inflação acabam sendo de fundamental importância para os gestores, uma vez que importantes decisões de investimento estão intimamente ligadas a elas.

Contudo, um aspecto que dificulta a previsão do comportamento inflacionário é que ele tende a se dar de forma não linear e até mesmo caótica, tornando difícil a sua correta estimação (RUGE-MURCIA, 2000; MALLIARIS, 2006). Essa característica do fenômeno pode tornar imprecisos os modelos mais simples de previsão – acessíveis às pequenas organizações –, uma vez que muitos deles necessitam de grandes manipulações de dados e/ou *softwares* especializados, incompatíveis com a rotina das organizações de menor porte. Assim, fica evidente a necessidade de encontrar modelos acessíveis e eficazes na previsão da problemática inflacionária para auxiliar os gestores de pequenas organizações na tomada de decisões de investimentos.

Em decorrência dessa problemática, o objetivo do presente esforço de pesquisa é verificar se determinados modelos de redes neurais artificiais são capazes de auxiliar na previsão da inflação em cenários complexos, tendo por base a realidade de organizações de pequeno porte.

## 2 AS REDES NEURAIS ARTIFICIAIS

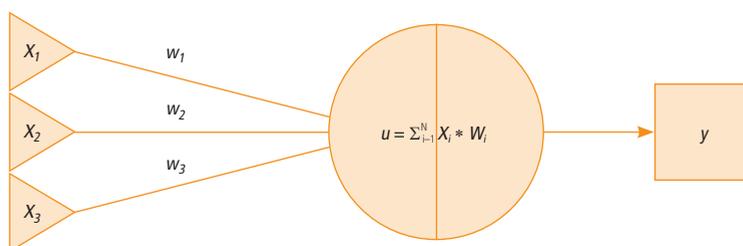
As redes neurais artificiais (RNA) são um ramo da inteligência artificial. Esses sistemas computacionais são inspirados nos neurônios biológicos e têm

como objetivo a solução de problemas complexos, a partir de um conjunto de variáveis independentes, como o reconhecimento de padrões e aproximações de funções (PRINCIPE; EULIANO; LEFEBVRE, 1999).

A origem da inteligência artificial se deu em 1943, quando McCulloch e Pitts desenvolveram o modelo matemático do primeiro neurônio artificial. Em 1957, Frank Rosenblatt desenvolveu o *Perceptron*, que funciona a partir de conceitos baseados no funcionamento dos neurônios orgânicos. No *Perceptron*, os dendritos são representados por entradas  $X_i$ , com pesos  $w_n$ , que são atribuídos a cada uma das conexões sinápticas, por meio de um processo de autoajustes. Já o resultado de saída  $Y$  é dado a partir de uma junção aditiva, aplicada a uma função de ativação  $\varphi(u)$ , que pode ser – dentre outras – a função sinal (KHANNA, 1990; PRINCIPE; EULIANO; LEFEBVRE, 1999). O esquema básico de um *Perceptron* pode ser visto na Figura 1.

FIGURA 1

ESQUEMA BÁSICO DO PERCEPTRON



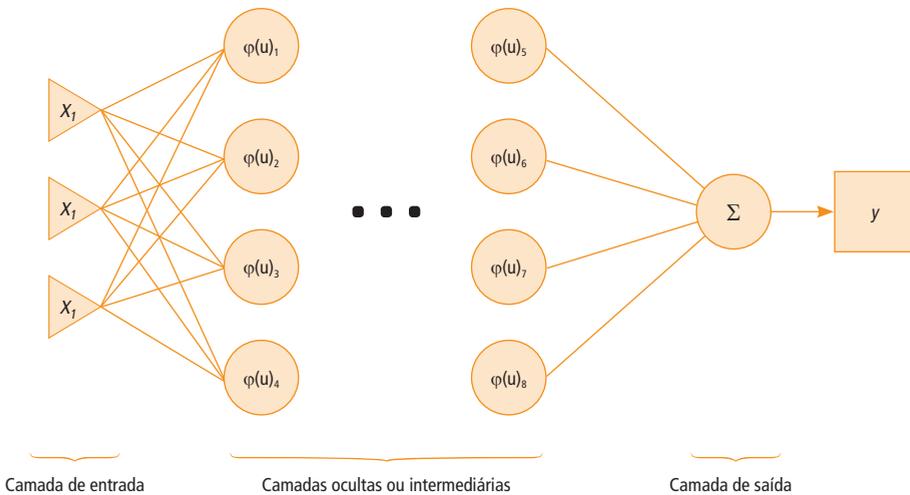
Fonte: Adaptada de Principe, Euliano e Lefebvre (1999, p. 103).

Fica evidente, no modelo, que as características dos *Perceptrons* os tornam incapazes de solucionar, individualmente, problemas que não sejam linearmente separáveis. Tal tipo de problema exigiria o uso de uma arquitetura estruturada por múltiplas camadas de *Perceptrons* (KHANNA, 1990). Porém, não havia uma solução para o treinamento de redes multicamadas, já que os algoritmos clássicos, utilizados para o treinamento de *Perceptrons* individuais, não resolviam o problema dessas arquiteturas de rede (PRINCIPE; EULIANO; LEFEBVRE, 1999). Tal restrição foi rompida por Rumelhart, Hinton e Williams (1986), por meio da proposição do algoritmo *Back-propagation*. Esse algoritmo tornou viável a implementação de arquiteturas de redes neurais formadas por múltiplas camadas de *Perceptrons*, denominadas, a partir de então, de *Multi Layer Perceptron* (MLP), cujo *layout* básico pode ser visto na Figura 2 (PRINCIPE; EULIANO; LEFEBVRE, 1999).

Fica evidente, em sua arquitetura, que as MLP representam uma aproximação de função complexa, obtida por meio de polinômios simples, cuja vantagem sobre um *Perceptron* individual é que tais redes são capazes de resolver problemas que não sejam linearmente separáveis. Ou seja, ao se autoajustarem por meio de um algoritmo de treinamento apropriado, as MLP podem produzir aproximações de funções capazes de fornecer uma boa capacidade de predição para ambientes mutáveis, regidos por comportamentos não lineares e relacionados a um grande volume de variáveis (HAYKIN, 1999; PRINCIPE; EULIANO; LEFEBVRE, 1999).

FIGURA 2

REDE MULTI LAYER PERCEPTRON

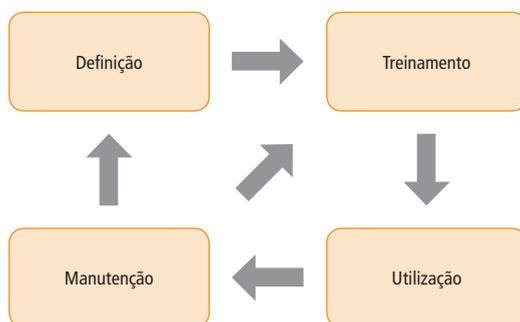


Fonte: Adaptada de Principe, Euliano e Lefebvre (1999, p. 227).

É importante ressaltar que, quando submetidas a ambientes em constante mudança, as MLP precisam receber manutenções sistemáticas, visando atualizar os pesos de suas conexões sinápticas às novas realidades impostas pelo meio e assim manter a capacidade de predição das MLP. O ciclo que vai do treinamento ao uso deve incluir revisões amplas, inclusive da própria definição da arquitetura da rede, conforme pode ser visto na Figura 3 (TAFNER; XEREZ; RODRIGUES FILHO, 1995).

FIGURA 3

### CICLO DE MANUTENÇÃO DAS REDES NEURAIIS ARTIFICIAIS



Fonte: Adaptada de Tafner, Xerez e Rodrigues Filho (1995).

Por sua capacidade de lidar com problemas não lineares em ambientes complexos, as redes neurais MLP parecem apropriadas para a solução de problemas, como a previsão da taxa da inflação, pois estão aptas a trabalhar com a não linearidade, que, segundo Ruge-Murcia (2000), assola a problemática inflacionária.

## 3 MÉTODO DE PESQUISA

Para execução deste trabalho, o treinamento das RNA propostas contou com um conjunto de dados de periodicidade mensal, coletados entre 1º de janeiro de 1998 e 31 de julho de 2008. As variáveis independentes foram selecionadas a partir da abordagem de Bresser-Pereira e Nakano (1984), enquanto a variável dependente do modelo foi a própria inflação do período.

O trabalho de Bresser-Pereira e Nakano (1984) se baseou em um levantamento sobre os principais fatores aceleradores, sancionadores e mantenedores da inflação. Esse levantamento foi feito a partir de uma confrontação entre as abordagens neoclássicas e monetaristas acerca da questão inflacionária. Com base na investigação e na revisão dos pressupostos básicos dessas escolas, os autores propuseram um avanço científico significativo para o entendimento dessa problemática, com a inclusão da inflação inercial como componente descritivo *adicional* aos modelos estudados (BRESSER-PEREIRA, 1996). O esforço desses pesquisadores brasileiros decifrou os questionamentos acerca da persistência inflacionária e complementou os modelos keynesianos e monetaristas, ao romperem com o pressuposto de que a inflação parte sempre de uma tendência zero.

Essa nova perspectiva é o sustentáculo dos modelos contemporâneos acerca da problemática inflacionária, que unificam os modelos anteriores e adicionam a própria inflação como um componente fundamental na compreensão da dinâmica inflacionária (BRESSER-PEREIRA; NAKANO, 1984; BRESSER-PEREIRA, 1996). A proposta de Bresser-Pereira e Nakano (1984) se tornou uma referência para as políticas de controle inflacionário e converge, inclusive, com a abordagem adotada contemporaneamente pelo Banco Central do Brasil. Por causa da importância dessa construção teórica, ela se constitui em uma alternativa adequada para ser o ponto de partida do presente esforço de pesquisa, uma vez que consolida as principais perspectivas presentes da área.

Os resultados do trabalho de Bresser-Pereira e Nakano (1984) encontram-se sintetizados no Quadro 1.

**QUADRO 1**

**FATORES ACELERADORES, MANTENEDORES E SANCIONADORES DA INFLAÇÃO**

FATORES ACELERADORES E SANCIONADORES DA INFLAÇÃO, SEGUNDO BRESSER-PEREIRA E NAKANO (1984)

- |                                    |                           |
|------------------------------------|---------------------------|
| • Taxa de juros básica da economia | • Base monetária restrita |
| • Taxa de câmbio                   | • Crescimento do PIB      |
| • Salários                         | • Inflação anterior       |
| • Resultado primário               |                           |

Fonte: Elaborado com base na abordagem de Bresser-Pereira e Nakano (1984).

A inflação de períodos anteriores foi tratada como variável independente, pois, segundo o modelo de Bresser-Pereira e Nakano (1984), a inércia inflacionária pode impactar na problemática inflacionária, promovendo a continuidade das taxas de inflação. Como índice de inflação, foi utilizado o IGP-DI, consolidado pela Fundação Getulio Vargas (FGV). O IGP-DI consistiu na escolha desse esforço de pesquisa, por ser um índice de inflação amplo, de fácil acesso e tradicionalmente usado em contratos adotados pelas pequenas e médias organizações.

A variável “emissão de moeda” foi utilizada, pois, nos períodos que antecederam o Plano Real, os monetaristas tratavam a emissão de moeda, para financiar o pagamento da dívida pública, como uma das principais causas da inflação. Essa variável foi representada pela consolidação das médias mensais da base monetária restrita, fornecidas pelo Banco Central do Brasil, cuja metodologia consiste na soma dos valores emitidos, adicionados das reservas bancárias, ou seja, é a consolidação de toda a disponibilidade de moeda no país (LOPES; ROSSETI, 2002).

Já a taxa de juros básica da economia é a mais conhecida e, ao mesmo tempo, a mais controversa das variáveis usadas pela política monetária do Banco Central do Brasil. Segundo Simonsen (1985), a taxa de juros básica tem um impacto similar ao dos depósitos compulsórios. O aumento em cascata dos juros na economia provoca a redução da oferta de crédito para o mercado, por causa do elevado custo de capital, tornando-a especialmente relacionada com a problemática inflacionária. Portanto, trata-se de uma variável essencial na criação de modelos que pretendem prever a inflação. Como fonte de dados para essa variável, foi utilizada a taxa Selic ponderada mensal, fornecida pelo Banco Central do Brasil.

No presente esforço de pesquisa, a taxa de câmbio também foi utilizada por causa de seu forte impacto inflacionário. A âncora cambial foi, inclusive, uma das políticas de contenção da inflação adotadas pelos formuladores do Plano Real. Tal política consiste na prática do uso de reservas cambiais para sustentar o câmbio em patamares baixos, mesmo diante da emissão de moeda. Assim, a própria queima de reservas acaba servindo de lastro para a moeda emitida, reduzindo a pressão inflacionária por meio da transferência desta para os valores praticados no exterior. Destarte, o câmbio atua de forma que, no limite, a inflação poderia ser direcionada para o valor da inflação no exterior (BATISTA JR., 1993). Como dado dessa variável, foi utilizado o valor médio mensal do dólar comercial, fornecido pelo Banco Central do Brasil.

Os salários médios também foram utilizados como variável independente do presente esforço de pesquisa, já que os salários têm uma característica particular de provocar a inflação por meio do conflito distributivo. Ao provocar aumentos de salários, a sociedade força o aumento da inflação ou a redução das margens dos capitalistas. Nesse sentido, espera-se que o capitalista não queira ver suas margens reduzidas, e, assim, o conflito distributivo caminha para um efeito em cascata, gerando inflação (BRESSER-PEREIRA; NAKANO, 1984). Como fonte de dados dos salários foi usado o salário médio mensal dos setores público e privado, consolidado pelo Dieese.

O resultado primário do governo também foi agregado como uma variável independente, já que afeta a necessidade de financiamento do Estado, exigindo ações por meio da aquisição de dívidas ou emissão de papel-moeda que culminam, em última instância, com a geração de inflação (BRESSER-PEREIRA; NAKANO, 1984). A fonte de dados para essa variável foi a Secretaria do Tesouro Nacional.

Por fim, o PIB se constitui na última variável independente adotada pelo presente esforço de pesquisa, já que ele se estabelece no próprio crescimento econômico do país e serve como referência da pressão produtiva deste. O PIB teve como fonte de dados o IBGE, a partir da nova metodologia adotada em 2007.

Uma vez que se pressupõe, com base em pesquisas prévias, que tais variáveis independentes são fatores causadores de inflação e, portanto, têm precedência sobre a inflação, essas variáveis foram defasadas em um, seis e doze períodos, em relação à inflação-alvo do treinamento. Com isso, pretendeu-se analisar a capacidade de predição dos modelos testados para períodos futuros de até um ano.

As variáveis apresentadas também têm significativas variações em suas ordens de grandeza. Assim, o uso dessas variáveis na grandeza original tornaria inviável a escolha de uma grandeza apropriada para a taxa de aprendizado das RNAs. Visando solucionar essa problemática do treinamento, os dados foram padronizados para reduzir as disparidades entre os valores nominais das variáveis independentes. Contudo, a variável dependente (inflação) foi mantida não padronizada, visando qualificar a rede a oferecer como resultado a própria inflação, sem a necessidade de reversão da padronização adotada.

A padronização utilizada segue a função 1, onde  $z$  é o valor padronizado;  $X$ , a variável de entrada;  $\bar{X}$ , a média da amostra; e  $S$ , o desvio padrão amostral. É importante notar que a padronização também exige uma manutenção constante da rede, uma vez que, a cada adição de novos dados ao conjunto de treinamento, obtêm-se uma nova média e um novo desvio padrão para a amostra do conjunto de treinamento, requerendo, assim, uma nova padronização dos dados propostos.

- Função 1: Padronização dos dados:

$$z = \frac{X - \bar{X}}{S}$$

Os impactos das observações atípicas sobre o treinamento da rede também foram considerados. Para tal, avaliou-se a alternativa de exclusão dos *outliers*, a partir de detecção multivariada. Porém, Hair et al. (1998, p. 66) afirmam que as observações atípicas “devem ser retidas a não ser que haja provas demonstrando que elas são verdadeiramente discrepantes e não são representativas de qualquer observação na população”<sup>1</sup>. De acordo ainda com Hair et al. (1998, p. 66), “se essas [ocorrências] representam um segmento da população, elas precisam ser mantidas para assegurar a capacidade de generalização [...]”<sup>2</sup> do modelo. Como os dados são todos representativos da população, uma vez que se constituem em uma consolidação da própria problemática inflacionária, e

<sup>1</sup> Do inglês: “*should be retained unless there is demonstrable proof that they are truly aberrant and not representative of any observations in the population*”.

<sup>2</sup> Do inglês: “*if they do represent a segment of the population, they should be retained to ensure generalizability [...]*”.

a capacidade de generalização do modelo é um elemento-chave para os objetivos do presente esforço de pesquisa, então não foram realizadas remoções das observações atípicas.

Foram escolhidas três arquiteturas de redes neurais MLP para avaliação, todas com o propósito de obter uma aproximação de função polinomial. Os modelos foram compostos por um *Perceptron* de saída, dotado de uma função de ativação linear  $\varphi_1(u)$  e de uma camada oculta com a função de ativação tangente hiperbólica  $\varphi_t(u)$ .

**QUADRO 2**

**FUNÇÕES DE ATIVAÇÃO USADAS NA AVALIAÇÃO DOS MODELOS**

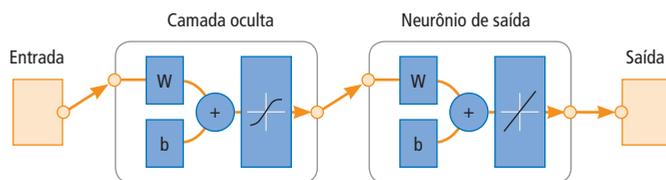
CAMADA DE SAÍDA	CAMADAS OCULTAS	
FUNÇÃO LINEAR	FUNÇÃO SIGMOIDAL	FUNÇÃO TANGENTE HIPERBÓLICA
$\varphi_1(u) = u$	$\varphi_s(u) = \frac{1}{1 + e^{-u}}$	$\varphi_t(u) = \frac{1 - e^{-u}}{1 + e^{-u}}$

Fonte: Adaptado de Principe, Euliano e Lefebvre (1999).

Realizaram-se testes preliminares também com a função sigmoial  $\varphi_s(u)$ , porém os resultados da função tangente hiperbólica  $\varphi_t(u)$  se mostraram mais consistentes. Assim, optou-se pela conformação das arquiteturas com a função que obteve os melhores resultados, produzindo modelos gerais, conforme ilustrado pela Figura 4.

**FIGURA 4**

**ARQUITETURA BÁSICA DA REDE**



Fonte: Elaborada com base em Matlab (2008a).

O número de *Perceptrons* ( $P$ ) presentes na camada oculta das redes foi determinado a partir de três abordagens tradicionais, dadas pelas funções que constam do Quadro 3, onde  $N$  é o número de variáveis independentes:

**QUADRO 3**

**NÚMEROS DE NEURÔNIO NA CAMADA OCULTA**

MODELO 1	MODELO 2	MODELO 3
$P(N) = N + 1$	$P(N) = 2N + 1$	$P(N) = \frac{N + 1}{2}$

Fonte: Elaborado com base em Principe, Euliano e Lefebvre (1999), Haykin (1999), Tafner, Xerez e Rodrigues Filho (1995) e experiências prévias dos autores.

**3.1 VALIDAÇÃO E TESTE DOS MODELOS**

Para validação e teste das redes, separaram-se aleatoriamente 30% das unidades estatísticas do conjunto de dados. Esse novo conjunto de dados foi utilizado, exclusivamente, para mensurar a capacidade de generalização de cada uma das redes. Além da avaliação do percentual de erro médio e do coeficiente de determinação ( $R^2$ ), em relação aos dados de controle, estes também foram submetidos à rede para a confrontação dos resultados pelo teste de comparação de médias de Wilcoxon.

O percentual de erro médio foi dado pela média simples do erro em relação ao resultado esperado, conforme a função 2.

- Função 2: Percentual de erro médio do modelo

$$\bar{E} = \frac{\sum_{i=1}^N \left( \frac{|X_r - X_p|}{X_r} \right)}{N}$$

onde  $\bar{E}$  é o percentual de erro médio;  $X_r$ ; o valor real da observação;  $X_p$ , o valor previsto pela rede; e N; o número de itens da amostra de teste.

O coeficiente de determinação foi utilizado para mensurar o quanto do fenômeno é explicado pela arquitetura e o treinamento em questão. Assim, determinou-se previamente que redes com coeficiente de determinação  $R^2$  menor do que 0,5 seriam classificadas como tendo baixa capacidade de explicar o fenômeno, redes com coeficiente de determinação  $R^2$  entre 0,5 e 0,7 seriam classificadas como tendo razoável capacidade de explicar o fenômeno, e redes com coeficiente de determinação  $R^2$  maiores do que 0,7 seriam tratadas como tendo boa capacidade de explicar o fenômeno.

O teste de Wilcoxon foi utilizado para verificar se existem diferenças entre as médias dos grupos de dados relacionados, sendo apropriado para situações em que é possível estabelecer um grupo de testes e um grupo de controle, como no caso das redes avaliadas (STEVENSON, 1986). É importante notar que o teste de Wilcoxon analisa as diferenças entre as médias amostrais, portanto pode ocorrer de uma rede ter um percentual de erro médio significativo para a previsão mensal, porém essa mesma RNA pode apresentar um resultado médio similar aos dados reais.

Para o teste de Wilcoxon, a hipótese nula ( $H_0$ ) adotada no transcórre deste trabalho foi de que a média dos dados obtidos pela rede neural é igual à média dos dados reais, e a hipótese alternativa ( $H_1$ ) foi de que a média dos dados obtidos pela rede neural é diferente da média obtida pelos dados reais. Tendo sido considerados aptos os modelos que confirmaram a hipótese nula ( $H_0$ ) com 95% de nível de confiança.

### 3.2 TREINAMENTO

O treinamento dos modelos foi feito através do pacote Neural Networks Tools do *software* Matlab 2008a. A escolha dos algoritmos de treinamento se deu visando à possibilidade de seu uso em programas de fácil acesso às pequenas organizações. Assim foi selecionado para o treinamento o algoritmo *back-propagation*, pois ele pode ser facilmente implementado em planilhas eletrônicas, inclusive em programas livres, mantendo-se, assim, o objetivo inicial de oferecer formas de previsão acessíveis às pequenas organizações.

O treinamento foi feito com gradiente descendente de  $10^{-10}$  e *momentum* de 0,9. Para todos os treinamentos, os pesos iniciais das conexões sinápticas foram definidos aleatoriamente. Testaram-se seis taxas de aprendizado entre 0,0001 e 0,05. Não foram usadas taxas de aprendizado com grandezas menores, pois elas exigem um maior número de interações e, assim, um maior número de linhas na construção das planilhas. Tais taxas de aprendizado exigiriam, portanto, maior poder computacional, elemento nem sempre disponível em organizações de pequeno porte. A taxa de aprendizado que apresentou o melhor desempenho foi 0,0001, como era de se esperar.

Para evitar a perda de capacidade de generalização da rede, por causa da alimentação, os treinamentos foram interrompidos sempre que houve 50 interações consecutivas, que promovessem persistência ou aumento do erro quadrático médio da rede, em relação ao conjunto de dados de validação. Ainda visando manter a usabilidade da rede em ambientes com recursos limitados,

realizaram-se treinamentos de, no máximo, 100 mil interações, pois se trata de um número de interações que pode ser realizado em períodos de tempo relativamente curtos, por meio de processadores de fácil aquisição, como o Intel Core 2 duo T8300 – modelo de processador utilizado para execução do presente trabalho. Foram feitos 30 treinamentos para cada modelo, com as regras de interrupção citadas anteriormente, sendo escolhido o de melhor desempenho para a análise comparativa.

## 4 RESULTADOS

A Tabela 1 apresenta os resultados do coeficiente de determinação  $R^2$  para os melhores treinamentos obtidos para cada um dos modelos de RNA avaliados e suas respectivas defasagens.

TABELA 1

### COEFICIENTES DE DETERMINAÇÃO $R^2$ DOS MELHORES TREINAMENTOS DE CADA MODELO

RESULTADOS DO COEFICIENTE $R^2$			
	MODELO 1	MODELO 2	MODELO 3
Defasagem de 1 mês	0,58	0,49	0,68
Defasagem de 6 meses	0,11	0,23	0,05
Defasagem de 12 meses	0,17	0,05	0,07

Fonte: Elaborada pelos autores.

O coeficiente de determinação demonstrou que o modelo 2 e os modelos de RNA com defasagem de seis e doze períodos não explicam o fenômeno da inflação de maneira ao menos razoável. A incapacidade de previsão futura em modelos não lineares normalmente se deve à perda de memória do fenômeno, oriunda da não inclusão de variáveis que podem não apresentar impactos significativos de curto prazo, mas podem apresentar impactos relevantes sobre o fenômeno no longo prazo (PRIGOGINE, 2002).

Já os modelos 1 e 3, com defasagem de um mês, apresentaram razoável capacidade de prever o fenômeno no curto prazo. Principalmente o modelo 3 que foi proposto pelos autores deste trabalho a partir de experiências prévias. O

modelo 3 atingiu um  $R^2$  de 0,68, a partir das variáveis defasadas de 1 mês, após um treinamento de 100 mil interações que produziu as conexões sinápticas apresentadas na Tabela 2.

**TABELA 2**

**CONEXÕES SINÁPTICAS DO TREINAMENTO DO MODELO 3  
PARA DEFASAGEM DE 1 PERÍODO**

CONEXÕES SINÁPTICAS DO TREINAMENTO DO MODELO 3 PARA DEFASAGEM DE 1 PERÍODO								
CONEXÕES SINÁPTICAS DOS NEURÔNIOS DA CAMADA OCULTA								
	VAR1	VAR2	VAR3	VAR4	VAR5	VAR6	VAR7	BIAS
N1	0,69317	0,29455	1,0101	1,0091	-0,1715	0,3437	0,39703	-0,01382
N2	0,48495	-0,52834	1,0315	0,16844	0,44471	-0,52253	0,48465	0,046456
N3	-0,74325	0,099977	-0,40745	0,10554	0,85678	0,80159	0,99551	0,33799
N4	0,46863	-1,1126	0,83024	-0,50573	0,46828	0,81485	-0,7983	-0,44085
CONEXÕES SINÁPTICAS DO NEURÔNIO DE SAÍDA								
	N1		N2		N3		N4	
NS	-1,7021		-0,97386		0,050077		1,5065	
BIAS	-0,084157							

Fonte: Elaborada pelos autores a partir do *software* Matlab (2008a).

O percentual de erro médio mostrou melhores resultados para os modelos 2 e 3, para defasagens de 1 mês. Os erros significativos, encontrados nas defasagens mais longas, confirmam a má capacidade de previsão dos três modelos nessa condição. A comparação entre as previsões com defasagem de 1 mês do modelo 3, que apresentou o melhor  $R^2$ , e os dados reais do conjunto de validação pode ser vista no Gráfico 1. Já os melhores resultados no percentual de erro médio de cada um dos modelos podem ser vistos na Tabela 3.

**TABELA 3**

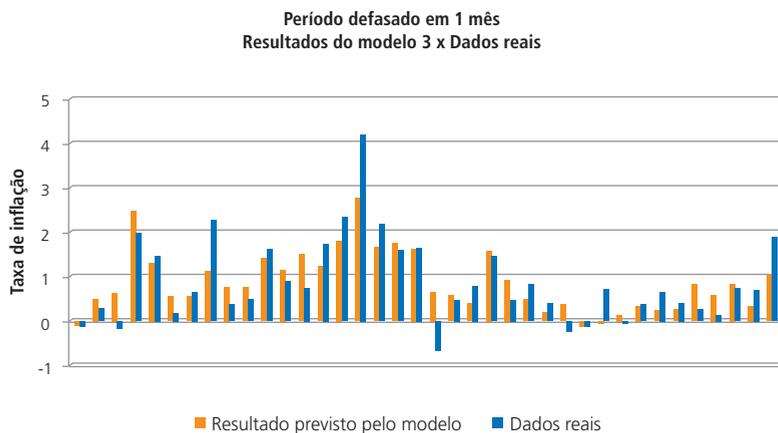
**PERCENTUAL DE ERRO MÉDIO DOS MELHORES TREINAMENTOS DE CADA MODELO**

RESULTADOS DO ERRO MÉDIO			
	MODELO 1	MODELO 2	MODELO 3
Defasagem de 1 mês	112%	75%	84%
Defasagem de 6 meses	243%	347%	292%
Defasagem de 12 meses	195%	324%	1135%

Fonte: Elaborada pelos autores.

**GRÁFICO I**

**RESULTADOS DO MODELO 3 VERSUS DADOS REAIS**



Fonte: Elaborado pelos autores.

O elevado percentual de erro médio do modelo 3, associado a um bom ajuste do  $R^2$ , pode ser explicado pela presença de algumas observações atípicas no conjunto de validação. No Gráfico 1, observa-se uma boa proximidade em um grupo considerável de observações, mas também algumas observações bastante discrepantes, que podem produzir significativas distorções na avaliação do percentual de erro médio.

O teste de Wilcoxon apontou que, salvo o modelo 3 associado a uma defasagem de 12 meses, todos os demais confirmaram a hipótese nula  $H_0$  com um nível de confiança de 95%. O melhor nível de significância coube ao modelo 3,

com defasagem de 1 mês (0,960). Os modelos que confirmaram  $H_0$  apresentaram uma previsão da média inflacionária igual à média inflacionária real. Dessa forma, o uso de múltiplas previsões futuras pode render um resultado médio para a inflação compatível com a média inflacionária prevista pelos modelos de RNA que confirmaram a hipótese nula.

## 5 CONSIDERAÇÕES FINAIS

Dentre os modelos testados, o mais apropriado para uso foi o modelo 3, que, além de ser capaz de produzir previsões com uma média próxima à real, conseguiu um desempenho razoável na mensuração de resultados mês a mês. Portanto, ele pode ser usado para atender, com razoável precisão, a essa expectativa. O modelo 1 também demonstra que as variáveis, popularmente aceitas, podem, sim, fornecer uma base satisfatória para os modelos de previsão mensal, baseados em redes neurais artificiais.

Finalmente, pode-se concluir que, partindo das variáveis tradicionalmente aceitas para esse tipo de mensuração, é possível utilizar as redes neurais artificiais para previsão média da inflação. Apesar das ressalvas referentes à presença de possíveis *outliers* e certa imprecisão nas previsões mês a mês, existem evidências de que o resultado médio, obtido pelos modelos testados, é capaz de produzir uma previsão similar à média de inflação real. Cabe também uma sugestão para trabalhos futuros, visando à avaliação de outras arquiteturas de redes neurais, para identificar aquelas que consigam apresentar maior capacidade de predição, inclusive com a inclusão de variáveis adicionais, que sejam acessíveis para o público em geral.

## ARTIFICIAL NEURAL NETWORKS IN INFLATION PREDICTION: APPLICATION LIKE ANALYSIS TOOL FOR FINANCIAL DECISIONS AT SMALL ORGANIZATIONS

### ABSTRACT

Estimation of inflation rates is crucial for managers, because investment decisions are closely linked to them. However, the behavior of inflation tends to be nonlinear and even chaotic, making it difficult to be estimated. This characteristic

may become simplest models, accessible for small organizations, inaccurate to forecasting the phenomenon, since many of these require large data manipulations and/or specialized software. This article aims to evaluate, through formal statistical analysis, the effectiveness of artificial neural networks in inflation forecasting at small organizations reality. ANNs are appropriated tools to measure the phenomena of inflation, as they are approximations of polynomial functions, capable of dealing with nonlinear phenomena. This article selected three basic models of *Multi Layer Perceptron* artificial neural networks, simple enough to be implemented whit open source spreadsheets. These three models were tested from a set of independent variables suggested by Bresser-Pereira and Nakano (1984), with a lag of one, six and twelve months. For that were used Wilcoxon test, coefficient of determination  $R^2$  and the average percent error of tested models. Data set was divided into two, one group used for artificial neural networks training and another group used to verify models predictive ability and their ability to generalize. This work concluded that certain models of artificial neural networks have a reasonable ability to predict inflation in the short run and constitute a reasonable alternative for this type of measurement.

## KEYWORDS

Inflation; Artificial neural networks; *Perceptron*; Small organizations; Analysis decisions.

## REDES NEURONALES ARTIFICIALES EN EL PRONÓSTICO DE LA INFLACIÓN: LA APLICACIÓN COMO UNA HERRAMIENTA PARA APOYAR EL ANÁLISIS DE LAS DECISIONES FINANCIERAS EN ORGANIZACIONES PEQUEÑAS

## RESUMEN

Las estimaciones de las tasas de inflación son cruciales para directivos ya que las decisiones de inversión están estrechamente vinculadas a ellas. Sin embargo, el comportamiento de la inflación tiende a ser no lineal e incluso caótico, lo que dificulta la correcta estimación. Esta característica del fenómeno puede hacer inexactos los modelos más simples para el pronóstico, accesible a las pequeñas organizaciones, ya que muchos de estos requieren manipulaciones de datos grandes

y/o *software* especializado. Este artículo tiene como objetivo evaluar, mediante un análisis estadístico formal, la eficacia de las redes neuronales artificiales en el pronóstico de la inflación, dentro de la realidad de las pequeñas organizaciones. Las RNA son herramientas adecuadas para medir el fenómeno de la inflación, ya que son aproximaciones de funciones polinómicas, capaz de hacer frente a fenómenos no lineales. Para este proceso fueron seleccionados tres modelos básicos de las redes neuronales artificiales *Multi Layer Perceptron*, aplicables desde la fuente de las hojas de cálculo de código abierto. Los tres modelos fueron probados a partir de un conjunto de variables independientes sugerido por Bresser-Pereira y Nakano (1984), con retraso de un, seis y doce meses. Para este fin, se utilizaron las pruebas de Wilcoxon, coeficiente de determinación  $R^2$  y el porcentaje medio de error de los modelos. El conjunto de datos se dividió en dos, siendo uno de los grupos utilizado para la formación de redes neuronales artificiales, mientras que otro grupo se utilizó para verificar la capacidad predictiva de los modelos y su capacidad de generalización. Con esto, el trabajo concluyó que ciertos modelos de redes neuronales artificiales tienen una capacidad razonable para predecir la inflación en el corto plazo y constituyen una alternativa razonable para este tipo de medición.

## PALABRAS CLAVE

Inflación; Redes neuronales artificiales; *Perceptron*; Organizaciones pequeñas; Análisis de decisiones.

## REFERÊNCIAS

- ASSAF NETO, A. *Finanças corporativas e valor*. São Paulo: Atlas, 2003.
- BATISTA JR., P. N. Dolarização, âncora cambial e reservas internacionais. *Revista de Economia Política*, v. 13, n. 3, p. 5-20, jul./set. 1993.
- BRESSER-PEREIRA, L. C. A inflação decifrada. *Revista de economia política*, v. 16, n. 4, p. 20-35, 1996.
- BRESSER-PEREIRA, L. C.; NAKANO, Y. Fatores aceleradores, mantenedores e sancionadores da inflação. *Revista de Economia Política*, v. 4, n. 1, p. 5-21, jan./mar. 1984.
- DAMODARAN, A. *Corporate finance: theory and practice*. New York: John Wiley & Sons, 1997.
- HAIR, J. F. et al. *Multivariate data analysis*. 5. ed. Upper Saddle River: Prentice Hall, 1998.
- HAYKIN, S. S. *Neural networks: a comprehensive foundation*. 2. ed. Upper Saddle River: Prentice Hall, 1999.
- JUSTER, F. T.; WACHTEL, P. Inflation and the consumer. *Brookings Papers on Economic Activity*, v. 1, n. 1, p. 71-121, 1972.

- KHANNA, T. *Foundations of neural networks*. Menlo Park: Addison-Wesley, 1990.
- LOPES, J. D. C.; ROSSETI, J. P. *Economia monetária*. 8. ed. São Paulo: Atlas, 2002.
- MALLIARIS, A. G. US inflation and commodity prices: analytical and empirical issues. *Journal of Macroeconomics*, v. 28, n. 1, p. 267-271, 2006.
- PRIGOGINE, I. *As leis do caos*. Tradução R. L. Ferreira. São Paulo: Editora Unesp, 2002.
- PRINCIPE, J. C.; EULIANO, N. R.; LEFEBVRE, W. C. *Neural and adaptive systems: fundamentals through simulations*. New York: John Wiley & Sons, 1999.
- RUGE-MURCIA, F. J. Uncovering financial markets' beliefs about inflation targets. *Journal of Applied Econometrics*, v. 15, n. 5, p. 483-512, 2000.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, v. 323, p. 533-536, 1986.
- SIMONSEN, M. H. A inflação brasileira: lições e perspectivas. *Revista de Economia Política*, v. 5, n. 4, p. 15-30, 1985.
- STEVENSON, W. J. *Estatística aplicada a administração*. São Paulo: Harpa, 1986.
- TAFNER, M. A.; XEREZ, M. D.; RODRIGUES FILHO, I. W. *Redes neurais artificiais: introdução e princípios de neurocomputação*. Blumenau: Eko, 1995.
- VAN HORNE, J. C.; WACHOWICZ JR., J. M. *Fundamentals of financial management*. 20. ed. Upper Saddle River: Prentice Hall, 1998.