

# Aplicação de um Modelo *Ensemble* de Redes Neurais Artificiais para Previsão de Séries Temporais não Estacionárias

Cristiano Leite de Castro<sup>1</sup>, Antônio de Pádua Braga<sup>2</sup>, Alessandro Vivas Andrade<sup>1</sup>

<sup>1</sup>Centro Universitário de Belo Horizonte - UNI-BH  
Av. Prof. Mário Werneck, 1685 - Estoril, 30455-610 Belo Horizonte, MG

<sup>2</sup>Universidade Federal de Minas Gerais, CPDEE - Campus da UFMG (Pampulha),  
Caixa Postal 209, 30161-970 Belo Horizonte, MG

ccastro@acad.unibh.br,

apbraga@cpdee.ufmg.br,

vivas@acad.unibh.br

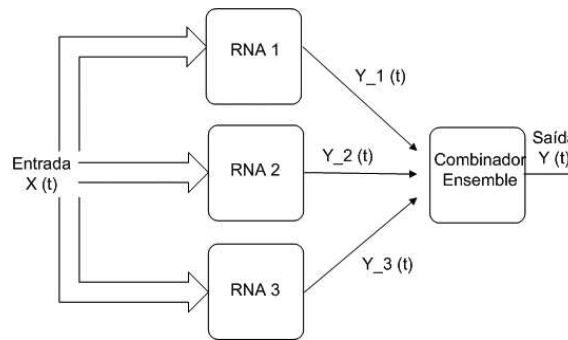
**Abstract.** *On forecasting of time series, better results have been obtained from combination of different forecasting models than from the choice of that single specific model with the best individual performance. This paper describes the use of the Ensemble method for the attainment of a forecasting model with high capacity of generalization. In this method, the outputs of different forecasting models are combined linearly to produce a global output. The Ensemble method was applied for the combination of Multi-Layer Perceptron nets that supply distinct solutions for the forecasting of a nonstationary series. Each net was trained differently by means of the manipulation of training parameters.*

**Resumo.** *Em previsão de séries temporais, melhores resultados têm sido obtidos pela combinação de diferentes modelos de previsão ao invés da seleção daquele que apresenta melhor desempenho individual. Este artigo descreve o uso do Método Ensemble para a obtenção de um modelo de previsão com alta capacidade de generalização. Neste método, as saídas de diferentes modelos de previsão são combinadas linearmente para produzir uma saída global. O método Ensemble foi aplicado para a combinação de redes Perceptrons de Múltiplas Camadas que fornecem soluções distintas para previsão de uma série não estacionária. Cada rede foi treinada diferentemente a partir da manipulação de parâmetros de treinamento.*

## 1. Introdução

Redes Neurais Artificiais (RNA's) [Braga et al., 2000] vêm sendo aplicadas extensivamente à previsão de séries temporais nos últimos anos [Refenes et al., 1997], [Kuan and White, 1994] e [Zhang, 2003]. Características inerentes aos modelos neurais, tais como aprendizado e generalização a partir de um conjunto de dados e universalidade na aproximação de funções contínuas multi-variáveis lineares e não lineares [Cybenko, 1988] fazem das RNA's uma ferramenta bastante atrativa para a tarefa de modelagem e previsão de séries não estacionárias.

As RNA's são capazes de aprender com os dados da série sem que haja conhecimento a priori a cerca do problema em estudo. Elas aprendem de exemplos e capturam os relacionamentos funcionais entre os dados, mesmo se tais relacionamentos são desconhecidos ou difíceis de descrever. Dessa forma, as RNA's são indicadas para problemas cujas soluções requerem conhecimentos que sejam difíceis de especificar, mas para os quais existam dados ou observações suficientes [Zhang et al., 1998]. Esta capacidade de aprender com os próprios dados da série (modelagem empírica) pode apresen-



**Figura 1: Combinação de três componentes produzindo diferentes soluções para uma mesma tarefa.**

tar problemas, pois nem sempre os relacionamentos entre as amostras são evidentes e as observações são frequentemente mascaradas por ruído.

Segundo [Sharkey, 1999], para a previsão de séries temporais, os melhores resultados são obtidos pela combinação de diferentes modelos de previsão ao invés da seleção daquele que apresenta melhor desempenho individual. A combinação de especialistas é conhecida na literatura como Máquina de Comitê [Haykin, 2001]. Esta máquina é capaz de agregar o conhecimento adquirido por especialistas para chegar a uma decisão global que é supostamente superior àquela alcançável por qualquer um deles atuando isoladamente.

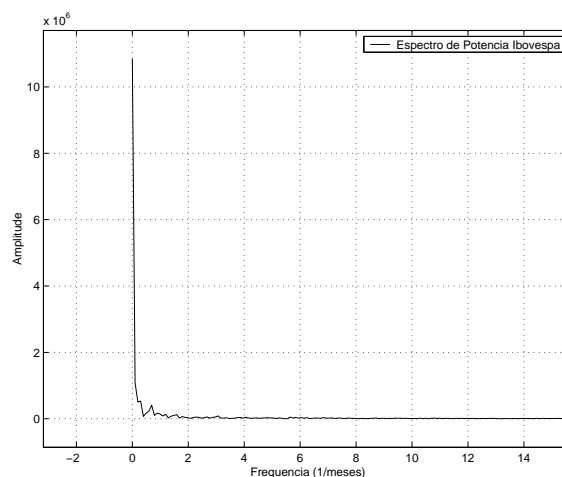
A principal motivação na combinação de modelos de previsão baseados em RNA's é a melhoria da habilidade de generalização [Sharkey, 1999]. Além disso, a combinação de diferentes RNA's permite que o modelo combinado minimize as falhas que possam ocorrer nas redes individuais.

Este artigo descreve o uso do método *Ensemble* [Haykin, 2001] para a obtenção de um modelo de previsão com alta capacidade de generalização. Neste método, as saídas de diferentes modelos de previsão são combinadas linearmente para produzir uma saída global. A Figura 1 ilustra a aplicação do Método *Ensemble* para a combinação de três componentes produzindo soluções diferentes para uma mesma tarefa. Os componentes são combinados para produzir uma única solução com alta capacidade de generalização.

O método *Ensemble* aqui descrito foi aplicado para a combinação de três redes Perceptrons de Múltiplas Camadas (MLP) distintas destinadas à previsão da série de índices de fechamento diário da bolsa de valores de São Paulo (*Ibovespa*).

Desde que a principal razão para a combinação de redes é melhorar seu desempenho, não há vantagem nenhuma em obter uma combinação composta por um conjunto de redes idênticas, ou seja, que generalizem de uma mesma maneira. A idéia é combinar redes que constituam soluções diferentes obtendo diferentes padrões de erro quando apresentadas a um mesmo conjunto de teste [Sharkey, 1999].

Existe um número de parâmetros de treinamento que pode ser manipulado para a obtenção de tais redes, como por exemplo: condições iniciais, dados de treinamento, topologia das redes e o algoritmo de treinamento. Para a previsão da série *Ibovespa* foram construídas redes MLP que obtinham soluções diferentes a partir do pré-processamento dos dados de treinamento: para a primeira rede MLP foi utilizada a série de resíduos  $Y_t$ , para a segunda, a série *Ibovespa* original com valores absolutos  $X_t$  e para a terceira, a série de retorno nominal  $Z_t$ . Além disso, foram utilizados quantidades diferentes de neurônios computacionais nas camadas intermediárias, variáveis de entradas distintas e diferentes algoritmos de treinamento. Tais procedimentos permitiram a obtenção de redes



**Figura 2: Espectrograma da série *Ibovespa*.**

MLP com diferentes capacidades de generalização.

A Seção 2 do artigo ilustra os procedimentos adotados para a seleção das variáveis de entrada dos modelos de previsão baseados em redes MLP. Na Seção 3 são apresentados os três modelos de previsão para a série *Ibovespa* e suas características individuais. A Seção 4 mostra como as soluções obtidas pelos modelos de previsão individuais foram combinadas para gerar um modelo de previsão com alta capacidade de generalização. A Seção 5 faz uma análise dos resultados obtidos e a Seção 6 traz as conclusões do trabalho.

## 2. Seleção das Variáveis de Entrada

A seleção das variáveis de entrada para as redes MLP destinadas à previsão da série *Ibovespa* foi realizada a partir de procedimentos comumente utilizados em análise de séries temporais [Brockwell and Davis, 1996].

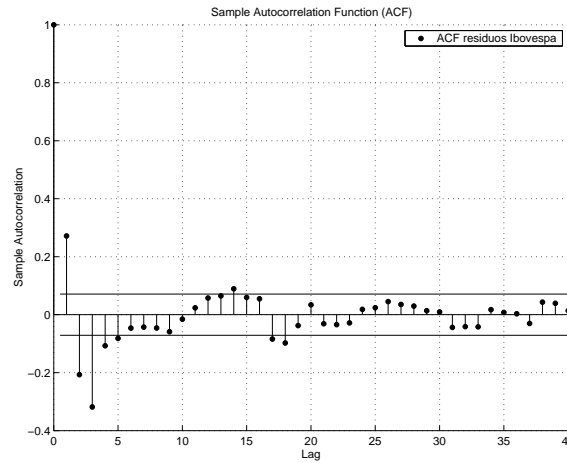
Foram consideradas séries de mercado internacional que representam os índices de fechamento diário das bolsas de valores de Tóquio (*Nikkei*), Londres (*FTSE100*), Nova Iorque (*Dow Jones*) e Frankfurt (*Dax30*) de 2 de janeiro de 2000 a 17 de março de 2003. Foram consideradas também séries de mercado nacional, como o fechamento da cotação diária *Dólar x Real* e o valor de fechamento do *Volume Negociado* (milhões de reais) diariamente na Bovespa no mesmo período.

### 2.1. Análise da Dependência através da Função de Autocorrelação Amostral

Inicialmente, uma análise gráfica foi realizada na série *Ibovespa* com o objetivo de identificar componentes determinísticos presentes na série, como ciclos, tendências e sazonalidades.

Analisando-se o gráfico da função espectral [Brockwell and Davis, 1996] da série, ilustrado pela Figura 2, o componente que representa a tendência foi detectado apresentando um pico de maior amplitude em uma região com frequência próxima de zero, o que corresponde a um período infinito, caracterizando um comportamento ao longo de todo o nível da série.

O componente de tendência  $m_t$  da série *Ibovespa*, foi estimado a partir da aplicação de um filtro finito de médias móveis bilateral (*two-sided*) com janela de tamanho 5 ( $q = 2$ ) conforme descrito pela Equação 1 [Brockwell and Davis, 1996],



**Figura 3: Autocorrelação da série ( $Y_t$ ) *Ibovespa*.**

$$m_t = (2q + 1)^{-1} \sum_{j=-q}^q X_{t-j}, \quad q + 1 \leq t \leq n - q \quad (1)$$

onde  $m_t$  representa o componente de tendência estimado,  $2q + 1$  representa o tamanho da janela para a aplicação do filtro,  $X_t$  representa a série temporal original e  $n$  representa o número de amostras da série. [Brockwell and Davis, 1996].

Após a estimação de  $m_t$ , os resíduos estacionários em média  $Y_t = X_t - m_t$  foram obtidos, para a série *Ibovespa*. Em seguida, realizou-se uma análise da dependência temporal da série *Ibovespa* para a identificação de atrasos relevantes.

As estimativas dos coeficientes de autocorrelação para a série de resíduos *Ibovespa* são apresentados na Figura 3. Observa-se pela Figura 3, a existência de coeficientes de autocorrelação fora do intervalo de confiança representado pelos limites de  $\pm 1.96\sqrt{n}$  no gráfico.

Conclui-se, portanto, que a série  $Y_t$  *Ibovespa*, no período de tempo considerado (janeiro de 2000 a março de 2003), apresenta um comportamento que depende linearmente de seus valores passados. Esta característica, apontada pela função de autocorrelação amostral ilustrada na Figura 3, sugere valores de entrada para os modelos de previsão baseados em Redes Neurais Artificiais [Braga et al., 2000], ou seja, pode-se utilizar valores passados da série  $Y_t$  *Ibovespa* para prever os valores seguintes.

## 2.2. Análise das Relações de Dependência Temporal entre as Séries de Mercado

Nesta fase, as séries *Nikkei*, *Dow Jones*, *Ftse100*, *Dax30*, *Ibovespa*, *Dólar x Real* e *Volume Negociado* foram consideradas como um processo multi-variável representando um vetor multi-dimensional  $(X_{t1}, X_{t2}, X_{t3}, X_{t4}, X_{t5}, X_{t6}, X_{t7})^T$ . Tal aproximação permitiu levar em consideração possíveis relações de dependência entre as séries componentes do processo multi-variável.

Em seguida, as funções de correlação cruzada [Brockwell and Davis, 1996] entre as séries *Nikkei*, *Dow Jones*, *Ftse100*, *Dax30*, *Dólar x Real*, *Volume Negociado* e *Ibovespa* foram estimadas. Os valores dos coeficientes de correlação cruzada estimados para os atrasos 0 e +1 são dados na Tabela 1.

Pelos valores dos coeficientes de correlação cruzada estimados para o atraso 0, pôde-se concluir que existem correlações significativas entre as séries *Ibovespa* e *Ftse100*, entre *Ibovespa* e *Dax30* e especialmente entre *Ibovespa* e *Dow Jones*. Essas séries

**Tabela 1: Coeficientes de correlação cruzada estimados para as séries de mercado.**

Séries de Mercado	Correlação no atraso 0	Correlação no atraso +1
Ibovespa e Dow Jones	0.44	0.19
Ibovespa e Ftse100	0.33	0.26
Ibovespa e Dax30	0.40	0.28
Ibovespa e Nikkei	0.06	0.33
Ibovespa e Dólar x Real	-0.40	-0.11
Ibovespa e Volume Negociado	0.17	0.05

apresentam-se correlacionadas no mesmo dia, refletindo as diferenças de horários de fechamento dos mercados nas cidades de Nova Iorque, São Paulo, Londres e Frankfurt, causadas pelos fuso horários. Sugeriu-se então que os mercados que fecham mais cedo possam servir de auxílio para a previsão dos mercados que fecham no mesmo dia, porém em horário posterior. Por exemplo, os mercados de Londres e Frankfurt podem servir como entrada para a construção de modelos de previsão para o mercado de São Paulo.

A Tabela 1 indica também que as séries *Ibovespa* e *Dólar x Real* apresentam-se correlacionadas. Isto pode ser observado para o atraso 0 cujo coeficiente é negativo. Sugere-se dessa forma que o mercado de câmbio e o mercado acionário funcionam simultaneamente durante o dia de forma antagônica, sendo que, na maioria das vezes uma queda na cotação diária do dólar está atrelada a uma alta do índice *Ibovespa* e vice versa. Verifica-se que a série *Dólar x Real* pode ser utilizada como variável de entrada para previsão do índice *Ibovespa* no mesmo dia, ou seja, no mesmo atraso relativo.

### 3. Modelos de Previsão com Redes Perceptrons de Múltiplas Camadas

Devido à sua capacidade de aproximar funções lineares e não lineares multi-variadas [Wang and Zhu, 2000], as redes neurais MLP (*Perceptrons de Múltiplas Camadas*) [Braga et al., 2000] foram utilizadas na obtenção de três modelos de previsão distintos para a série *Ibovespa*. Essa arquitetura de rede neural consiste em um conjunto de unidades sensoriais que constituem a entrada da rede, uma ou mais camadas ocultas de elementos computacionais (neurônios) e uma camada de saída de elementos computacionais. O sinal de entrada se propaga para a frente através da rede camada por camada [Haykin, 2001]. Nas Seções seguintes são apresentadas as características dos três modelos.

#### 3.1. Modelo de Previsão para a Série de Resíduos *Ibovespa*

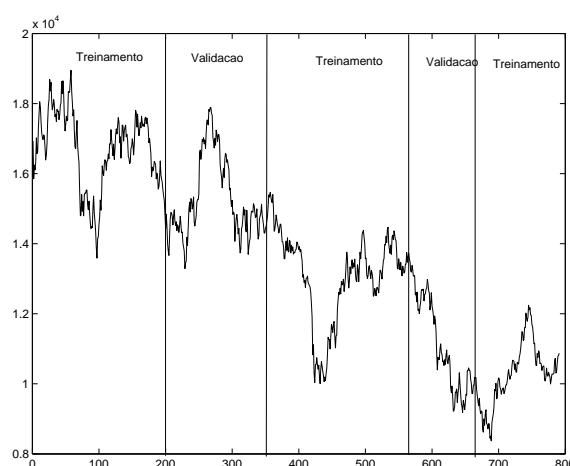
O primeiro modelo de previsão desenvolvido para a série *Ibovespa* utilizou a série de resíduos ( $Y_t$ ). Essa série foi obtida com a estimação e eliminação do componente de tendência ( $m_t$ ) a partir da série *Ibovespa* original ( $X_t$ ). A obtenção da série de resíduos ( $Y_t$ ) é descrita com detalhes na Seção 2.1.

O modelo de previsão para a série de resíduos ( $Y_t$ ) é baseado em uma rede MLP com topologia 6:25:1, ou seja, 6 entradas, 25 neurônios na camada intermediária com função de ativação do tipo tangente hiperbólica e um único neurônio na camada de saída com função de ativação do tipo linear.

A seleção das variáveis de entrada foi realizada a partir das análises estatísticas realizadas nas Seções 2.1 e 2.2. A Tabela 2 traz as variáveis (séries) de entrada selecionadas para o modelo.

**Tabela 2: Variáveis de Entrada dos Modelos de Previsão.**

Modelo de Previsão	Variáveis de Entrada
MLP para Previsão de $Y_t$	Dax30( $t$ ), Ftse100( $t$ ), Ibovespa( $t - 1, t - 2, t - 3$ ), Dólar x Real( $t$ )
MLP para Previsão de $X_t$	Dax30( $t$ ), Ftse100( $t$ ), Ibovespa( $t - 1, t - 2, t - 3$ ), Dólar x Real( $t$ )
MLP para Previsão de $Z_t$	DowJones( $t$ ), Dax30( $t$ ), Ftse100( $t, t - 1$ ), Ibovespa( $t - 1, t - 2, t - 3$ ), Dólar x Real( $t, t - 1$ )

**Figura 4: Faixas utilizadas pelos conjuntos de treinamento e validação do modelo.**

A rede MLP foi treinada com um total de 600 amostras de entrada (conjunto de treinamento) e o restante da série (191 amostras) foi utilizado no conjunto de validação. A verificação do desempenho alcançado pelo modelo foi realizada comparando os valores obtidos na saída da rede com a série de resíduos *Ibovespa* no instante  $t$  (conjunto de teste). O algoritmo de treinamento utilizado foi o Backpropagation [Hinton, 1981] associado ao método de controle de generalização conhecido como Validação Cruzada [Stone, 1987].

### 3.2. Modelo de Previsão para a Série de Valores Absolutos *Ibovespa*

O segundo modelo desenvolvido para a série *Ibovespa* realiza a previsão da série original de valores absolutos ( $X_t$ ). Ele é baseado em uma rede MLP com topologia 6:30:1 e conta com as mesmas variáveis de entrada utilizadas no modelo de previsão para a série de resíduos ( $Y_t$ ) apresentado na Seção 3.1. A Tabela 2 mostra as variáveis de entrada selecionadas.

A rede MLP foi treinada com um total de 541 pontos, representando 541 padrões de entrada, e o restante da série foi utilizado na validação do modelo. Os vetores de entrada da rede foram divididos em cinco faixas distintas, três para treinamento e duas para validação, conforme ilustrado pela Figura 4. Este procedimento permitiu que o mapeamento de entrada-saída realizado pela rede MLP para os dados do conjunto de validação, pudesse capturar toda a faixa de valores expressa pela série de índices original, *Ibovespa*( $t$ ).

O algoritmo de treinamento que obteve melhores resultados para esse modelo foi o algoritmo multi-objetivo (MOBJ) proposto por [Teixeira et al., 1999]. O desempenho

de previsão do modelo foi testado comparando os resultados obtidos na saída da rede com a série *Ibovespa* original de valores absolutos no instante  $t$ .

### 3.3. Modelo de Previsão para a Série de Retorno Nominal *Ibovespa*

O terceiro modelo realiza a previsão para a série de retorno nominal *Ibovespa* ( $Z_t$ ) obtida a partir da transformação nos dados da série *Ibovespa* original ( $X_t$ ) através da equação 2.

$$Z_t = 100 \times \frac{X_t - X_{t-1}}{X_{t-1}}, t = 1, \dots, N. \quad (2)$$

onde  $Z_t$  representa a série expressa através de suas variações percentuais diárias e  $X_t$  representa a série temporal original [Brockwell and Davis, 1996].

O modelo de previsão para a série de variações ( $Z_t$ ) conta com uma rede MLP com topologia 9:35:1. A Tabela 2 mostra as variáveis de entrada selecionadas.

A rede MLP foi treinada com um total de 600 amostras ou padrões de entrada e o restante da série (191 amostras) foi utilizado para a obtenção do conjunto de validação. A verificação do desempenho alcançado pelo modelo foi realizada comparando os valores obtidos na saída da rede com a série de retorno nominal *Ibovespa* no instante  $t$ . O algoritmo de treinamento que obteve os melhores resultados foi o algoritmo Levenberg Marquardt [Hagan and Menhaj, 1994].

A Tabela 3 sumariza as características dos três modelos baseados em redes MLP para previsão da série *Ibovespa*.

**Tabela 3: Características dos Modelos de Previsão.**

Modelo de Previsão	Dados de Treinamento	Topologia da Rede	Algoritmo
MLP 1	Série ( $Y_t$ )	6:25:1	Backpropagation
MLP 2	Série ( $X_t$ )	6:30:1	MOBJ
MLP 3	Série ( $Z_t$ )	9:35:1	Lev. Marquardt

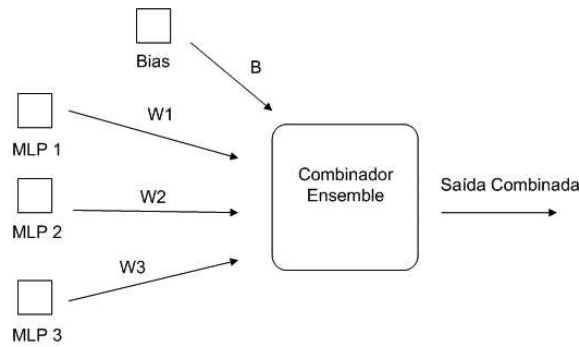
## 4. Aplicação do Método *Ensemble*

Após o treinamento das redes *Perceptrons de Múltiplas Camadas* (MLP) foi utilizado o Método *Ensemble* para combinação de suas saídas individuais [Haykin, 2001]. Este método foi aplicado a partir da construção de uma rede com um neurônio computacional de saída apresentando função de ativação linear. As variáveis de entrada da rede correspondem às respectivas saídas de cada modelo de previsão para a série *Ibovespa*, além de uma unidade sensorial com entrada +1 representando o termo de polarização (ou *bias*). A Figura 5 ilustra o Método *Ensemble* desenvolvido para a combinação dos três modelos de previsão para a série *Ibovespa*.

Os parâmetros do método *Ensemble* (pesos e *bias*) foram obtidos a partir da equação 3,

$$W = Y(X^T X)^{-1} X^T, \quad (3)$$

onde  $W$  é a matriz de parâmetros da rede contendo os pesos e *bias*,  $X$  é a matriz de padrões de entrada da rede,  $Y$  é o vetor de saída desejada da rede e  $(X^T X)^{-1} X^T$  é a pseudo-inversa de  $X$ , que minimiza a soma dos erros quadráticos  $\sum \|WX - Y\|^2$ , proporcionando solução ótima para a matriz de pesos  $W$  [Lipschutz, 1991].



**Figura 5: Método *Ensemble* aplicado à combinação dos modelos de previsão para a série *Ibovespa*.**

## 5. Resultados

Um grande número de medidas de desempenho podem ser encontradas na literatura de previsão de séries temporais e cada uma possui suas vantagens e limitações [Makridakis et al., 1982]. Neste artigo, a eficácia dos modelos de previsão desenvolvidos para a série *Ibovespa* foi medida, inicialmente, a partir da raiz dos erros quadráticos médios (*RMSE*) descrita pela equação 4.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{Y}_t - Y_t)^2}{N}}, \quad (4)$$

onde  $\hat{Y}_t$  é o valor de previsão no instante  $t$ ,  $Y_t$  é a saída desejada no instante  $t$  e  $N$  é o número de amostras para as quais os termos de erro são calculados.

A Tabela 4 ilustra as medidas de erro *RMSE* obtidas pelos modelos de previsão da série *Ibovespa*. Para fins de comparação entre os modelos de previsão desenvolvidos, a medida *RMSE* foi calculada para o conjunto, localizado entre as amostras 600 e 791 da série *Ibovespa*. Para todos os modelos, os erros de previsão foram calculados a partir da série de valores absolutos,  $X_t$ .

**Tabela 4: Medidas de erro (RMSE) para os modelos de previsão para a série *Ibovespa*.**

Modelos de Previsão	RMSE
MLP para Previsão de $Y_t$	70.40
MLP para Previsão de $X_t$	68.36
MLP para Previsão de $Z_t$	114.35
Modelo Combinado <i>Ensemble</i>	37.30

Observa-se pela Tabela 4 que o modelo combinado apresenta a menor medida de erro (RMSE) entre todos os modelos desenvolvidos, mostrando ser mais eficaz na previsão dos índices *Ibovespa*.

Em um segundo estágio, os modelos de previsão propostos são comparados com o método de previsão ingênuo (*naive*), onde  $y(t) = y(t-1)$ , ou seja, o valor do índice de fechamento a ser previsto é igual ao valor do índice no dia anterior [Braga et al., 2002]. Segundo [Zhang, 2003], a série *naive* é uma boa previsão para uma série financeira que assemelha-se a uma série do tipo passeio aleatório (*random walk*), como é o caso da série *Ibovespa*. Para mostrar que os modelos de previsão propostos obtiveram resultados melhores que o modelo *naive*, foi utilizada uma medida de razão  $r$  descrita pela equação 5 [Braga et al., 2002],



$$r = \frac{\sqrt{\sum_i (Y_i^p - Y_i)^2}}{\sqrt{\sum_i (Y_i^n - Y_i)^2}}, \quad (5)$$

onde  $Y^p$  é o valor previsto pelo modelo de previsão,  $Y$  é o valor da série de índices original (desejada) e  $Y^n$  é o valor previsto pelo modelo *naive*.

Esta razão é conhecida como *Theil's U* ou coeficiente de diferença [Gaynor and Kirkpatrick, 1995] e serve para medir a eficiência de um modelo de previsão. Para valores de  $r$  abaixo de 1, a equação 5 indica que o erro obtido pela rede neural MLP é menor que o erro obtido pelo predictor *naive*. Como pode ser observado pela Tabela 5, os modelos de previsão propostos são melhores que o modelo *naive* apresentando um valor inferior a 0.5.

**Tabela 5: Medida de razão obtida pela comparação dos modelos de previsão propostos com o modelo ingênuo.**

Modelos de Previsão	<i>Theil's U</i>
MLP para Previsão de $Y_t$	0.0990
MLP para Previsão de $X_t$	0.0935
MLP para Previsão de $Z_t$	0.2615
Modelo Combinado <i>Ensemble</i>	0.0028

Segundo [Gaynor and Kirkpatrick, 1995], modelos de previsão com valores de *Theil's U* iguais ou menores a 0.55 são considerados confiáveis. O valor apresentado pelo modelo de previsão combinado, vide Tabela 5, indica que esse modelo apresenta elevado grau de confiabilidade para a previsão dos fechamentos diários da bolsa de valores de São Paulo.

## 6. Conclusões

O artigo descreveu o uso do método *Ensemble* para a obtenção de um modelo de previsão com alta capacidade de generalização. O método adotado, foi utilizado para a combinação de três redes MLP que fornecem soluções distintas para previsão da série *Ibovespa*. Conforme observado na Seção 5, o modelo combinado obteve em média melhores resultados de previsão que cada rede componente em sua forma individual. Ele foi capaz de combinar a habilidade de previsão alcançada por cada um dos modelos, obtendo dessa forma, melhor capacidade de generalização que os modelos individuais.

Qualquer forma de combinação de RNA's deve ser realizada a partir de componentes individuais que alcancem diferentes padrões de erro quando apresentados a um mesmo conjunto de teste, ou seja, que generalizam a informação aprendida de forma diferente. Nas redes MLP individuais destinadas à previsão da série *Ibovespa* isto foi obtido com a seleção de diferentes parâmetros de treinamento para cada uma das redes. Os parâmetros manipulados foram: conjunto de treinamento, número de neurônios computacionais nas camadas intermediárias das redes e algoritmos de treinamento.

## 7. References

### Referências

Braga, A. P., Carvalho, A. C. P. L. F., Ludermir, T. B., Almeida, M. B., and Lacerda, E. (2002). *Radial Basis Function Networks In Modelling and Forecasting Financial*

- Data: techniques of non-linear dynamics*. Soofi, A. S and Cao, L. Editors - Kluwer Academic Publishers, Boston, USA.
- Braga, A. P., Lurdemir, T. B., and Carvalho, A. C. P. L. F. (2000). *Redes Neurais Artificiais: Teoria e Aplicações*. LTC - Livros Técnicos e Científicos Editora S.A., Belo Horizonte, Brasil.
- Brockwell, P. J. and Davis, R. A. (1996). *Introduction to Time Series and Forecasting*. Springer Verlag, New York, USA.
- Cybenko, G. (1988). Approximation by superpositions of a sigmoidal function. Technical report, University of Illinois, Urbana, IL, USA.
- Gaynor, P. E. and Kirkpatrick, R. C. (1995). *Introduction to time-series modeling and forecasting in business and economics*. McGraw-Hill, Appalachian State University, Boone, North Carolina, USA.
- Hagan, M. T. and Menhaj, M. (1994). Training feedforward networks with the marquardt algorithm. *IEEE Transactions on Neural Networks* vol.5 no.6 pgs.989-993.
- Haykin, S. (2001). *Redes Neurais - Princípios e Prática*. Bookman, Porto Alegre, Brasil.
- Hinton, G. (1981). Connectionist learning procedures. *Artificial Intelligence*, 40, pgs. 185-234.
- Kuan, C. M. and White, H. (1994). Artificial neural networks: an economic perspective. *Economic Reviews* 13(1), pgs. 1-91.
- Lipschutz, S. (1991). *Álgebra Linear*. Editora Makron Books, Sao Paulo, Brasil.
- Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibdon, M., and Lewandowski, R. (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of Forecasting* 1(2), pgs. 111-153.
- Refenes, A. P. N., Burgess, A. N., and Bentz, Y. (1997). Neural network in financial engineering: A study in methodology. *IEEE Transactions on Neural Networks*, Vol. 8, n. 6, pgs. 1222-1267.
- Sharkey, A. J. C. (1999). *Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems*. Springer Verlag London Limited, London, UK.
- Stone, M. (1987). Cross-validation: A review. *Mathematische Operationforschung Statistische, Serie Statistics* 9, pgs. 127-139.
- Teixeira, R. A., Braga, A. P., Takahashi, R. H. C., and Saldanha, R. R. (1999). Improving generalization of mlp with multi-objective optimization. *Neurocomputing*, 35(1-4), pgs. 189-194.
- Wang, Z. O. and Zhu, T. (2000). An efficient learning algorithm for improving generalization performance of radial basis function neural networks. *Neural Networks* 13(4-5), pgs. 545-553.
- Zhang, G. P. (2003). Time series forecasting using a hybrid arima and neural network model. *Neurocomputing* Vol. 50 pgs. 159-175.
- Zhang, G. P., Patuwo, B. E., and Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting* 14 pgs. 35-62.