

# Efeito da Janela Temporal em um Modelo de Rede Neural Artificial na Predição do IBOVESPA

R. V. Aguiar, I. R. S. Casella

e-mail: rafael.aguiar@aluno.ufabc.edu.br

e-mail: ivan.casella@ufabc.edu.br

**Abstract** - *The study of methods for time series prediction has a great value for social and economical development. Among these techniques, the bio-inspired models have showed great results in the time series prediction problem. Thereby, this article describes an analysis of the effect of the time window length in the performance of the bio-inspired non-linear model based on Artificial Neural Networks for predicting the behaviour of the most important Brazilian stock market index IBOVESPA considering different metrics.*

**Palavras-chave:** *Série Temporal, Janela Temporal, Rede Neural Artificial, IBOVESPA.*

## Introdução

É inerente ao processo de tomada de decisão, efetuar previsões acerca do futuro. Neste processo, comumente supõe-se que existe correlação entre os dados passados e futuros [1], ou seja, pode-se considerar que eventos observados até o presente momento podem influenciar aqueles que ainda não ocorreram. Neste sentido, pode-se citar como exemplo, o interesse de governos e empresas da iniciativa privada em avaliar o comportamento do mercado de ações [2]. A importância deste processo no âmbito governamental é evidenciada pelo fato do desempenho da economia nacional poder ser estimado pelo comportamento do mercado de ações. Se o mercado tende a uma queda contínua, é pertinente supor que um período de recessão pode estar próximo [3]. Já para uma empresa, o preço de suas ações reflete seu momento atual, bem como suas perspectivas futuras, se as ações tendem à alta, condiciona-se a empresa para realização de novos financiamentos de investimento.

Comumente, diz-se que os valores numéricos do indicador a ser predito é uma série temporal que representa uma função amostra de determinado processo estocástico subjacente [4]. Desta forma, tem-se que o problema de predição consiste em aproximar a equação dinâmica do processo estocástico em uma função que represente adequadamente os valores observados e que torne possível estimar os valores instantes futuros [1].

Para o processo de predição, é comum, por

exemplo, o uso dos seguintes modelos estatísticos [5]: (i) *Autoregressive* (AR), (ii) *Moving Average* (MA), (iii) *Autoregressive Moving Average* (ARMA) e (iv) *Autoregressive Integrated Moving Average* (ARIMA).

Além dos modelos estatísticos, também é possível empregar técnicas bio-inspiradas não-lineares oriundas da Inteligência Artificial (IA), como as Redes Neurais Artificiais (RNA) [6], para a tarefa de predição. As RNA pertencem ao segmento da IA denominado conexionista [7], em que supõe-se que o comportamento inteligente é de natureza fisiológica, armazenado nas conexões sinápticas entre unidades elementares do cérebro denominadas neurônios. Em uma RNA, é feita a substituição destes neurônios fisiológicos por elementos abstratos de *software* e/ou *hardware* e uma vez que determina-se a quantidade de neurônios interconectados e adaptados aos dados de entrada, a RNA exhibe capacidade de generalização, tornando-a capaz de fornecer respostas satisfatórias para estímulos ainda não conhecidos [6].

Neste contexto, o presente trabalho tem como objetivo a representação do IBOVESPA por uma série temporal estocástica, a modelagem dessa série por uma RNA do tipo Multilayer Perceptron (MLP) e a identificação e otimização dos parâmetros mais relevantes da RNA de modo a maximizar a precisão da predição, tendo como especial atenção o comprimento da janela temporal dos dados de entrada da RNA.

## Descrição do Modelo Proposto

A predição de valores futuros em uma série temporal, assume a existência de autocorrelação entre o valor de uma função  $s(\cdot)$  no instante atual  $n$  e seus valores passados para os instantes defasados  $n - 1, \dots, n - N$ , em que  $N$  representa a ordem do modelo de autoregressão [5]. Além disso, pode-se ter correlação cruzada entre a série temporal aproximada pela função  $s(\cdot)$  de interesse e uma série explicativa dada pela função  $e(\cdot)$  em seu instante atual  $m$  e seus valores defasados  $m - 1, \dots, m - M$  em que  $M$  representa a ordem do modelo de correlação cruzada [1].

A estrutura das funções  $s(\cdot)$  e  $e(\cdot)$  pode originar diferentes modelos, *e.g.* pode-se utilizar ambas as funções e assumir que elas são lineares, dando origem ao modelo ARMA exibido na Equação 1.

$$\hat{s}(n) = \sum_{i=1}^N a(i)s(n-i) + \sum_{j=1}^M b(j)e(m-j) \quad (1)$$

Sendo  $\hat{s}(\cdot)$  a aproximação de  $s(\cdot)$ ,  $a(i)$  o coeficiente de autoregressão linear para associado ao instante  $n-i$  e  $b(j)$  é o coeficiente de regressão cruzada associado ao instante  $m-j$ .

Apesar de útil e aplicável em diversas situações, o modelo ARMA apresenta algumas limitações, pois nem sempre a hipótese de linearidade é sustentada nas séries temporais de interesse [5]. Desta forma, torna-se relevante o uso de modelos não-lineares como o modelo bio-inspirado baseado em RNA, composto pela interconexão de neurônios computacionais [6], para aplicação ao problema de predição [7]. Conforme ilustrado na Figura 1, a topologia das RNA é comumente subdividida em três camadas [7]:

**Camada de entrada** - Coleta dados do ambiente externo e, eventualmente, aplica pré-processamento para adequá-los ao sistema;

**Camada oculta** - Determina a correlação entre as entradas e saídas observadas para, posteriormente, aplicá-la às entradas não observadas;

**Camada de saída** - Disponibiliza os resultados obtidos a partir do processamento realizado nas camadas anteriores;

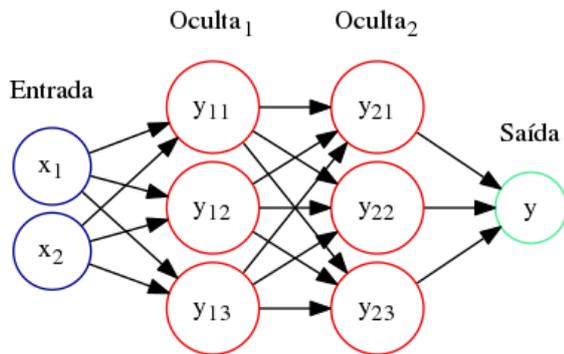


Figura 1 – Arquitetura de RNA em Camadas.

Em particular ao escopo deste trabalho, foi escolhida a RNA baseada na arquitetura MLP,

obtida pela interconexão de neurônios do tipo *perceptron*, por ela ser capaz de tratar problemas não linearmente separáveis [6].

A MLP é caracterizada pela Equação 2, onde o vetor de pesos sinápticos  $w_{ij}$  do neurônio  $i$  da camada  $j$ , é responsável por codificar o aprendizado do neurônio e a parcela denominada *bias*  $\lambda_{ij}$ , é necessária para deslocar a reta de separação do espaço de estados a partir da origem [6]. A seguir, efetua-se o produto interno entre o vetor de entrada  $\vec{x}$  e  $w_{ij}$ , gerando o resultado intermediário  $\theta_{ij}$  que é aplicado à função de ativação  $g(\cdot)$ , obtendo-se assim, a saída  $y_{ij}$  que é propagada aos neurônios das camadas à frente, até a saída final [7].

$$y_{ij} = g(\theta_{ij}) = g(\vec{x} \cdot \vec{w}_{ij} + \lambda_{ij}) \quad (2)$$

Desta análise, pode-se notar a importância do treinamento, em que os vetores de ponderação  $w_{ij}$  e *bias*  $\lambda_{ij}$  são ajustados a partir do algoritmo de aprendizagem selecionado [8]. Durante o treinamento, a RNA identifica as correlações presentes nos dados, possibilitando uma generalização do modelo para entradas ainda não observadas.

Existem diferentes modelos de treinamento, neste trabalho será utilizado o aprendizado supervisionado, realizado pelo algoritmo *Backpropagation* [8]. Este algoritmo caracteriza-se pela apresentação de pares ordenados  $(X, T)$  em que  $X$  é a entrada e  $T$  é a saída desejada correspondente. Para cada apresentação do par de treinamento  $(X, T)$ , efetua-se o ajuste dos pesos comparando a saída obtida com aquela desejada, resultando no sinal de erro a ser minimizado retroativamente [6].

Com base no exposto, no que concerne a aplicação do modelo RNA à predição de séries temporais, tem-se o seguinte fluxo:

1. Preparação dos dados - transformar o conjunto de dados em uma representação adequada à RNA;
2. Divisão dos dados - dividir o conjunto de dados em partições para treinamento e predição;
3. Treinamento - treinar a RNA com os dados previamente conhecidos, para ela representar um comportamento igual a série temporal em análise. Nesta etapa deve-se ter cautela com o fenômeno de *overfitting* em que a RNA adequa-se perfeitamente aos dados de treinamento, mas não generaliza, ou seja, não adequa-se bem aos dados de predição;
4. Predição - Avaliar o comportamento da RNA mediante sua aplicação para horizontes fu-

turos de predição, ou seja, fora do conjunto de treinamento.

Na etapa de treinamento, bem como ao final do desenvolvimento da RNA, é necessário avaliar sua adequação aos dados, dito de outra maneira, mensurar seu desempenho [6]. Entre as métricas de desempenho da RNA, destaca-se, o *Mean Squared Error* (MSE) e o *Akaike Information Criterion* (AIC), definidos nas equações 3 e 4, respectivamente. Em que  $s(\cdot)$  é a série temporal sob análise,  $\hat{s}(\cdot)$  é sua aproximação pelo modelo de RNA,  $K$  é o total de observações e  $p$  é o número de parâmetros da RNA.

$$MSE = \frac{1}{K} \sum_{k=0}^K [s(k) - \hat{s}(k)]^2 \quad (3)$$

$$AIC = \ln \left\{ \frac{1}{K} \sum_{k=0}^K [s(k) - \hat{s}(k)]^2 \right\} + \frac{2p}{K} \quad (4)$$

## Análise dos Resultados

Nesta seção, descreve-se a aplicação do modelo de RNA ao problema de predição do IBOVESPA. Inicialmente, considerou-se a predição de uma série temporal obtida a partir da Equação 5, que é amostrada com período  $T_a = 10$  ms, resultando na série temporal de tempo discreto  $P[n]$ , criada com o intuito de representar o comportamento aleatório com determinado grau de correlação entre amostras temporais do IBOVESPA nas análises iniciais, possibilitando a verificação da corretude do algoritmo implementado antes de sua validação com dados reais.

$$P(t) = 0,2X + \begin{cases} 2\text{sen}(4\pi t) + \cos(20\pi t), & t \leq 1 \\ 2\text{sen}(4\pi t) + t^2, & t > 1 \end{cases} \quad (5)$$

Esta equação é composta por uma parte determinística e outra estocástica. O componente determinístico exibe mudança de tendência após  $t > 1$ , já o componente estocástico é representado por um processo gaussiano com  $X \sim \mathcal{N}(0, 1)$ .

Com base no procedimento de quatro etapas definido anteriormente, realizou-se a subdivisão dos dados em 85% para treinamento e 15% para predição. Empregando o método de busca exaustiva, para diversas configurações obtidas mediante a variação do comprimento da janela de entrada ( $M$ ), número de camadas ocultas e o número de

neurônios por camada ( $H$ ); avaliou-se o desempenho da RNA utilizando as métricas anteriormente expostas: (i) MSE e (ii) AIC.

Para cada configuração, determinou-se o MSE e o AIC do conjunto de dados completos, *i.e.*, treinamento e predição. Ao final, a configuração otimizada obtida é composta por uma janela de 14 pontos e uma camada oculta com 30 neurônios. O resultado obtido com a MLP otimizada é exibido na Figura 2. Em que o MSE ótimo é  $9,13 \cdot 10^{-2}$  e seu AIC correspondente é 2,39.

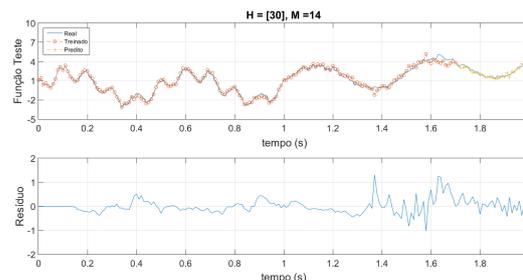


Figura 2 – Resultado da MLP Otimizada para a Série Temporal Definida pela Equação 5.

A análise da função de teste apresentou-se importante para entender o processo de representação de séries temporais mediante uma RNA, identificar seus parâmetros mais relevantes de desempenho e verificar a corretude do algoritmo implementado.

Após a análise dos resultados obtidos com a série temporal de teste, aplicou-se a RNA ao problema de predição do preço de fechamento do dia seguinte do IBOVESPA com base em seus valores passados, considerando o período de 01/01/2005 até 31/12/2015. Os dados utilizados estão disponíveis em [9].

Assim como no caso da função de teste, na análise do modelo de RNA aplicado ao problema de predição do IBOVESPA, adotou-se o método de busca exaustiva para realizar a otimização da configuração da RNA segundo o critério do menor MSE global, em que determinou-se a configuração de duas camadas ocultas, com 20 e 40 neurônios, respectivamente. Após a verificação da influência de cada parâmetro no desempenho global do sistema, apresenta-se na Tabela 1 alguns dos resultados obtidos para a configuração otimizada, considerando diferentes comprimentos da janela temporal.

A análise dos resultados apresentados indica que, enquanto o AIC exibe consistência, elevando-se conforme aumenta-se o comprimento da janela temporal, o MSE exibe comportamento os-

catatório. As oscilações observadas podem ser explicadas verificando que, para janelas pequenas, o modelo exibe baixa capacidade de memória, não adequando-se bem à correlação existente nos dados, ao passo que, para janelas de grande comprimento, o modelo torna-se sensível ao ruído presente nos dados. Desta forma, pode-se ter pontos de mínimos locais de MSE em função do comprimento da janela.

Tabela 1 – Desempenho para Diferentes Janelas.

Janela	MSE ( $10^5$ )	AIC
2	7,90	14,30
4	8,04	14,30
7	7,98	14,40
10	6,56	14,20
14	7,21	14,40

A partir da Tabela 1, tem-se que a configuração otimizada segundo o critério MSE é obtida empregando uma janela temporal de 10 amostras.

Após o processo de otimização da RNA empregando dados reais do IBOVESPA, obteve-se o MSE global otimizado igual a  $6,56 \cdot 10^5$  e seu respectivo AIC igual a 14,20. O comportamento da predição em conjunto com a comparação à série temporal real do IBOVESPA, bem como o sinal de erro quadrático instantâneo, é exibido na Figura 3.

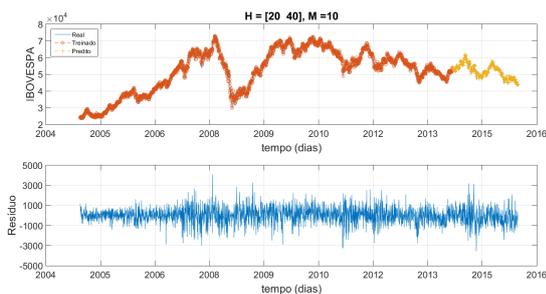


Figura 3 – Comportamento da MLP Otimizada para Dados Reais do IBOVESPA.

O valor  $6,56 \cdot 10^5$ , obtido para o MSE pode parecer elevado, entretanto, a variância da série temporal real é  $1,45 \cdot 10^8$  e a razão entre o MSE e a variância, denominado MSE normalizado, é  $4,29 \cdot 10^{-3}$ . Deste modo, o MSE é apenas 0,43% da variância da série temporal, indicando que o modelo é capaz de explicar satisfatoriamente os dados e reduzir significativamente o erro de predição.

## Conclusões

Neste trabalho, foi apresentado um modelo bio-inspirado na RNA baseada na arquitetura MLP para a predição dos valores futuros do IBOVESPA. A MLP foi otimizada empregando dados reais da bolsa de valores de São Paulo e os resultados obtidos mostraram que ela apresenta satisfatória precisão, mesmo na presença de variações abruptas e mudanças de tendência. As análises realizadas destacaram ainda a importância do correto dimensionamento da janela temporal na capacidade de predição da MLP.

## Referências

- [1] P. A. Morettin and C. M. C. Toloi, *Análise de Séries Temporais*. Edgard Blucher, 2 ed., 2006.
- [2] E. Fortuna, *Mercado Financeiro Produtos e Serviços*. Qualitymark, 4 ed., 2005.
- [3] A. Assaf, *Mercado Financeiro*. Atlas, 2 ed., 1999.
- [4] S. Kay., *Fundamentals of Statistical Signal Processing*. Prentice Hall, 1993.
- [5] D. C. Montgomery and C. L. Jennings, *Introduction to Time Series Analysis and Forecasting*. Wiley, 1 ed., 2008.
- [6] S. O. Haykin, *Neural Networks and Learning Machines*. Pearson, 3 ed., 2008.
- [7] A. de Padua Braga and A. P. Ludermir, *Redes Neurais Artificiais - Teoria e Prática*. LTC, 2 ed., 2011.
- [8] M. Khashei and M. Bijari, *An artificial neural network (p, d, q) model for timeseries forecasting*, vol. 37. Elsevier, 2010.
- [9] G. Finance, *Índice IBOVESPA*. ago 2017. <https://finance.google.com/finance/historical?startdate=Jan+01%2C+2005&enddate=Oct+22%2C+2017&num=200&q=INDEXBVMF:IBOV> [Acessado em: 22/10/2017].