

COMPARAÇÃO ENTRE ABORDAGENS DE PREVISÃO DE SÉRIES TEMPORAIS BASEADAS EM MODELOS HOLT-WINTERS E REDES NEURAIS ARTIFICIAIS

HEITOR ANDRE KIRSTEN¹, THIAGO RHOTEN² E LEANDO DOS SANTOS COELHO¹

¹*Pós-Graduação em Engenharia de Produção e Sistemas (PPGEPS), PUCPR*

²*Graduação em Engenharia Mecatrônica (Controle e Automação)*

Imaculada Conceição, 1155, CEP 80215-901, Curitiba, PR

*E-mails: heitorkirsten@gmail.com; thiago.rothen@gmail.com;
leandro.coelho@pucpr.br*

Abstract— The time series forecasting is a problem that has received special attention due its fundamental importance in analysis and decision making in all segments. The Holt-Winters is a kind of exponential forecasting model, known for dealing with time series elements of trend and seasonality. Artificial neural networks have been major tools in the application of time series forecasts, and its capacity for parallelism and nonlinearity draw attention to their use. This paper aimed to make the comparison between the forecasting models of Holt-Winters, MultiLayer Perceptron (MLP) neural network and Radial Basis Function (RBF) neural network, to make one-step-ahead forecasting of Nasdaq Composite Index (^IXIC) historical time series. The result found was that the MLP neural network model outperformed Holt-Winters and RBF neural network models in one-step-ahead forecasting of NASDAQ index time series.

Keywords— Forecasting, Time Series, Holt-Winters, Artificial Neural Networks

Resumo— A previsão de séries temporais é um problema que tem recebido especial atenção devida sua fundamental importância nas análises e tomadas de decisões em todos os segmentos. O Holt-Winters é um modelo de previsão exponencial, conhecido por lidar com elementos de tendência e sazonalidade de uma série temporal. As redes neurais artificiais têm se mostrado grandes ferramentas na aplicação de previsões de séries temporais, bem como sua capacidade de paralelismo e não-linearidade chamam a atenção para sua utilização. O objetivo deste artigo foi a comparação entre os modelos de previsão Holt-Winters, redes neurais MLP (*MultiLayer Perceptron*) e RBF (*Radial Basis Function*) quando da previsão um passo a frente para a série temporal financeira do índice Nasdaq Composite (^IXIC). O resultado encontrado foi que a rede neural MLP superou os modelos Holt-Winters e a rede neural RBF para a previsão um passo a frente da série temporal do índice NASDAQ Composite.

Palavras-chave— Previsão, Séries Temporais, Holt-Winters, Redes Neurais Artificiais

1 Introdução

Uma série temporal é uma coleção de observações feitas sequencialmente ao longo do tempo. A característica mais importante deste tipo de dado é que as observações vizinhas são dependentes umas das outras. Enquanto para modelos de regressão a ordem não importa, para as séries temporais essa ordem é crucial.

Freqüentemente existe uma diferença de tempo entre um evento ou necessidade iminente e a ocorrência desse evento ou necessidade. Esta diferença de tempo é a principal razão para prever e planejar. A previsão é uma importante ajuda para um planejamento eficiente e eficaz (MAKRIDAKIS *et al.*, 1998).

O fato de que a previsão de séries temporais é um problema que tem recebido especial atenção, dado que a previsão de séries temporais é fundamental em análises e tomada de decisões, como já constatado por Abelém (1994) e Morettin e Toloï (1987), chamou a atenção por suas dificuldades e complexidades.

A verdade é que vários trabalhos já estudaram a comparação entre modelos de previsão para uma série temporal. No entanto, a vasta quantidade de

possibilidades para fazer a previsão de uma série temporal e as várias conclusões encontradas pelos autores faz com que se justifique a pesquisa comparativa entre alguns dos modelos mais populares encontrados, como o Holt-Winters e as redes neurais MLP (*Multilayer Perceptron*) e RBF (*Radial Basis Function*).

Este artigo contribui no aspecto de ajudar a elucidar qual modelo de previsão pode se adequar melhor para as séries temporais financeiras. Acaba contribuindo também para as comparações dos modelos lineares com os modelos não-lineares de previsão. É um diferencial deste trabalho a comparação entre os modelos Holt-Winters e as redes neurais MLP e RBF para a previsão de séries financeiras.

O restante deste artigo foi organizado de forma. A seção 2 apresenta os fundamentos do modelo Holt-Winters. A seção 3 apresenta as redes neurais MLP e RBF. Nas seções 4 e 5, a série temporal estudada e os resultados das previsões, respectivamente, são detalhados. Por fim, a seção 6 apresenta a conclusão sobre as previsões efetuadas.

2 Modelo Holt-Winters

Corrar e Theóphilo (2004) explicam que Holt-Winters “é um método destinado ao tratamento de séries temporais que apresentam tendência e sazonalidade. Operando de forma semelhante ao método de Holt, o método de Winters incorpora parâmetros adicionais para o tratamento da sazonalidade da série”.

Resumindo, o modelo Holt-Winters, segundo Lewis (1997), assume que uma demanda baseada em tendência e sazonalidade é composta dos seguintes elementos: i) um elemento de nível, estacionário, que não exhibe tendência ou sazonalidade, mas que pode incluir alguma variação aleatória; ii) um elemento de tendência, em adição ao elemento de nível; e iii) um elemento de sazonalidade em adição ao elemento de tendência e nível.

Makridakis *et al.* (1998) apresenta as equações de Nível (1), Tendência (2), Sazonalidade (3) e Previsão (4) do modelo Holt-Winters, tal que:

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad (1)$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \quad (2)$$

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s} \quad (3)$$

$$F_{t+m} = (L_t + b_t m)S_{t-s+m} \quad (4)$$

onde L_t é o elemento nível para o período t , b_t o elemento tendência para o período t , S_t o elemento sazonalidade para o período t , F_{t+m} o valor da previsão para o período m , Y_t o valor observado no período t , s o intervalo de tempo da sazonalidade, m o número do período previsto, α a constante de ajuste do nível, β a constante de ajuste da tendência e γ a constante de ajuste da sazonalidade. Sendo que $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$ e $0 \leq \gamma \leq 1$.

No modelo Holt-Winters existe a necessidade de inicializar os valores de nível, tendência e sazonalidade para aplicar as fórmulas (1), (2), (3) e (4).

Sobre a inicialização dos valores de nível, tendência e sazonalidade, Makridakis *et al.* (1998) menciona que, para determinar os valores da sazonalidade é necessário pelo menos um período sazonal de dados históricos completo. Enquanto isso, os valores iniciais do nível e da tendência são iniciados no período s . O nível pode ser iniciado com a média da primeira época sazonal, conforme a equação (5).

$$L_s = \frac{1}{s}(Y_1 + Y_2 + \dots + Y_s). \quad (5)$$

Nota-se que o nível é uma média móvel de ordem s , com o intuito de retirar o efeito da sazonalidade. Para iniciar o elemento da tendência,

Makridakis *et al.* (1998) usa duas épocas sazonais, e expõe a equação (6).

$$b_s = \frac{1}{s} \left[\frac{Y_{s+1} - Y_1}{s} + \frac{Y_{s+2} - Y_2}{s} + \dots + \frac{Y_{s+s} - Y_s}{s} \right]. \quad (6)$$

Cada termo é uma estimativa da tendência sobre uma sazonalidade completa. Para o valor inicial da tendência faz-se uma média de todos os termos.

A seguir, encontra-se então o valor da sazonalidade para os primeiros períodos da época sazonal. Esse índice é encontrado dividindo-se os primeiros valores observados pelo valor inicializado do elemento nível, conforme a equação

$$S_1 = \frac{Y_1}{L_s}, S_2 = \frac{Y_2}{L_s}, \dots, S_s = \frac{Y_s}{L_s}. \quad (7)$$

Depois, para iniciar o primeiro elemento do nível deve-se dividir o valor observado no período s pelo índice sazonal do mesmo período s .

A ordem a ser seguida para encontrar os valores dos elementos para cada um dos períodos é a seguinte: i) calcular o elemento nível; ii) calcular o valor da tendência; e iii) calcular o índice da sazonalidade.

Quando todos os períodos estiverem com os valores preenchidos usa-se a equação (4) para determinar os valores projetados.

3 Redes Neurais Artificiais

As redes neurais artificiais, originalmente desenvolvidas para mimetizar sistemas neurais biológicos básicos (o cérebro humano, basicamente), são compostas por um número de elementos simples de processamento interconectados denominados neurônios (ZHANG *et al.*, 1998). Estes elementos simples, os neurônios, são, segundo Haykin (2001), implementados utilizando-se componentes eletrônicos ou através de simulação por programação em um computador digital.

3.1 Redes Neurais MLP

Segundo Haykin (2001), o *perceptron* é a forma mais simples de uma rede neural usada para a classificação de padrões denominados linearmente separáveis. O *perceptron* de uma única camada consiste de um único neurônio com pesos sinápticos e *bias*. Esse modelo de *perceptron* de um único neurônio é limitado a realizar classificação de padrões com apenas duas classes.

Para resolver o problema do *perceptron* de um único neurônio é necessário adicionar mais camadas na estrutura do *perceptron*, essas camadas adicionais são chamadas camadas ocultas.

A figura 1 mostra a estrutura básica de uma rede neural MLP, com uma camada de entrada, duas camadas ocultas e uma camada de saída. Esta mesma

estrutura pode ser representada na forma de equação, segundo Santos (2005), conforme a equação,

$$y = \varphi_s \left\{ \sum_{i=1}^m w_i \varphi_i \left(\sum_{j=1}^n w_{ij} \varphi_j + b_j \right) + b_s \right\}. \quad (8)$$

Santos (2005) define três aspectos básicos do projeto de uma rede neural MLP: i) a determinação do número de camadas ocultas; ii) a determinação do número de neurônio em cada uma das camadas; e iii) a especificação dos pesos sinápticos que interconectam os neurônios nas diferentes camadas da rede.

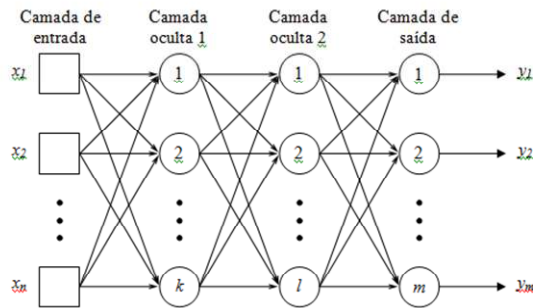


Figura 1 – Exemplo de estrutura de uma rede neural MLP

Segundo Haykin (2001), a camada oculta não faz parte da entrada ou saída da rede. A camada oculta e seus neurônios são responsáveis por fazer a rede aprender tarefas complexas extraindo progressivamente as características mais significativas dos padrões de entrada.

Segundo Guerra e Coelho (2002), o algoritmo Levenberg-Marquardt é uma alternativa viável e mais eficiente que o algoritmo de retropropagação do erro (*error backpropagation*) clássico, sendo esse o validado neste artigo.

3.2 Redes Neurais RBF

As redes neurais RBF são de uma classe de redes neurais com uma camada oculta, onde cada neurônio é ativado por uma função de base radial, de onde é derivado seu nome. Os neurônios na camada oculta fornecem um conjunto de funções não-lineares, de base radial, que constituem uma base arbitrária para os padrões de entrada (SANTOS, 2005).

Esse tipo de rede foi originalmente desenvolvido para interpolação de dados em espaços multidimensionais, conforme diz Castro (2001). Esse problema pode ser formulado pela fórmula $y_i = F(u_i)$, onde y_i é um conjunto de escalares e u_i é um conjunto de vetores. Desta forma, se a função $F()$ for definida analiticamente ela poderá mapear valores do vetor u_i que não pertenciam ao conjunto original de treinamento, no conjunto de escalares y_i . Para que a função $F(u_i)$ seja uma solução analítica ela pode ser definida da equação (9).

$$F(u) = \sum_{i=1}^n w_i \varphi(\|u - u_i\|^2), \quad (9)$$

onde w_i é um coeficiente (o peso da conexão neste caso) e $\varphi(\|u - u_i\|^2)$ representa uma função radial. Existindo n funções radiais, o centro de cada função é dado por u_i . Desta forma, $\|u - u_i\|$ representa a distância euclidiana entre o centro de cada função radial (u_i) e do vetor de entrada (u).

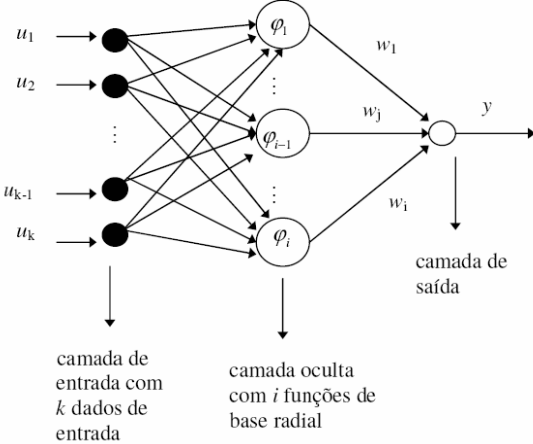


Figura 2 – Estrutura básica de uma rede neural RBF. Adaptado de Santos (2005).

Na sua forma mais básica as redes neurais RBF são compostas por três camadas: camada de entrada, camada oculta e camada de saída. A primeira camada, de entrada, tem o objetivo de receber os sinais do ambiente exterior e passá-los para a próxima camada. A segunda camada, a camada oculta, é formada por um conjunto de neurônios, sendo que cada neurônio é associado a um vetor de parâmetros chamado de centro. Cada neurônio calcula a distância Euclidiana entre o respectivo centro e o valor passado pela camada de entrada com base em uma função de base radial. Como a função de base radial é não-linear, como a função gaussiana, a camada oculta passa a ser não-linear. A terceira camada é a camada de saída, que é a resposta linear do mapeamento feito pela camada oculta. É um mapeamento linear dos valores não-lineares da camada oculta. A figura 2 mostra a estrutura básica de uma rede neural RBF (PAIVA, 1999).

A saída de uma rede neural RBF, usando uma função gaussiana na camada oculta, pode ser expressa na equação

$$y = \sum_{i=1}^n w_i \left(\exp\left(-\frac{\|u - u_i\|^2}{2\sigma^2}\right) \right), \quad (10)$$

onde w_i é o peso da sinapse entre a camada oculta e a camada de saída, $\exp\left(-\frac{\|u - u_i\|^2}{2\sigma^2}\right)$ é a função gaussiana presente na camada oculta e y representa a saída da rede neural RBF.

Conforme Santos (2005), os métodos mais comumente utilizados para o ajuste dos parâmetros livres são o *k-means* para o início e atualização dos centros e o mínimos quadrados para atualização dos pesos sinápticos.

4 Série Nasdaq

Segundo a Nasdaq (2009), o índice NASDAQ Composite compreende todas as ações comuns listadas na bolsa de valores NASDAQ, domésticas ou internacionais. O índice foi lançado no ano de 1971 com o valor base 100 e atualmente inclui mais de 3000 empresas, quase todas da área de tecnologia e biotecnologia. O índice representa a maior bolsa de valores eletrônica do mundo.

Para fazer a coleta dos dados utilizou-se a base de dados do site de finanças da Yahoo: www.finance.yahoo.com.

Foram coletados os dados relativos ao Índice NASDAQ Composite (^IXIC) a partir da data 01/01/1995 até a data 31/12/2008. A periodicidade dos dados é diária e o preço utilizado é o de fechamento. Os dias onde não houve pregão, como feriados e fins de semana, foram removidos do universo.

A figura 3 apresenta o gráfico da série temporal do índice Nasdaq utilizada neste trabalho.

Desta maneira o universo é compreendido pelos dias onde houve negociação entre o dia 01/01/1995 até o dia 31/12/2008, totalizando um montante de 3526 amostras.

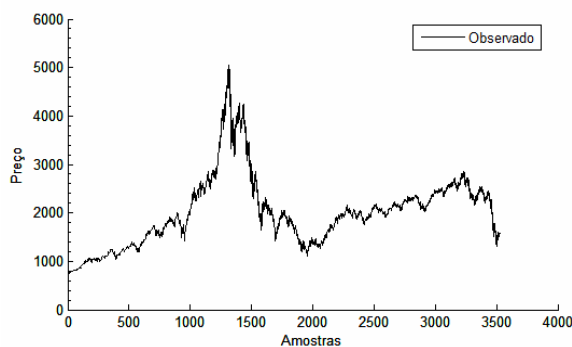


Figura 3 – Série temporal do índice Nasdaq Composite (^IXIC)

A série temporal financeira Nasdaq passou por uma normalização nos dados a fim de deixar os valores da série dentro do intervalo de 0 e 1. Procedimento necessário para ser possível a utilização de modelos de Redes Neurais na previsão. Para tanto a série passou por uma transformação onde todos os valores da série foram divididos pelo valor mais alto da série.

Também houve uma segmentação na série onde 80% dos dados da amostra foram utilizados para estimação do modelo, 10% para teste (a fim de encontrar o menor MSE (*Mean Squared Error*)) e os 10% restantes para a validação do modelo.

5 Resultados da previsão

A fim de comparar os resultados dos modelos de previsão para a série temporal do índice Nasdaq, foram selecionados três índices normalmente utilizados: MSE, MAPE (*Mean Absolute Percentage Error*) e R^2 (coeficiente de determinação).

5.1 Previsão utilizando Holt-Winters

Para a previsão da série Nasdaq usando Holt-Winters foram estimadas sazonalidades de tamanho 1 até 400, sendo que para cada sazonalidade estimada os parâmetros α , β e γ foram testados independentemente com valores entre 0 e 1 a um passo de 0,1 a cada etapa.

Considerando o MSE como índice de desempenho da seção de teste e buscando o menor valor para este índice, a melhor sazonalidade encontrada foi de 10, enquanto os parâmetros α , β e γ encontrados foram 0,9, 0 e 0,3.

Visto que os valores estimados na seção de estimação foram iguais aos valores observados da mesma seção, os índices de desempenho para esta seção não são considerados.

A figura 4 apresenta o erro quadrático da previsão usando o modelo Holt-Winters.

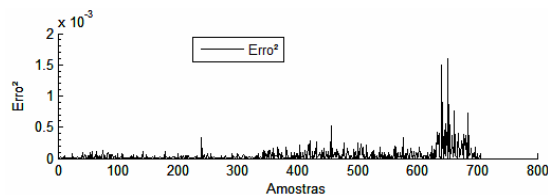


Figura 4 – Erro quadrático da previsão da série Nasdaq usando o modelo Holt-Winters para os segmentos de teste e validação

Na seção de teste os valores do MSE, MAPE e R^2 foram, respectivamente, 0,000017796, 0,6899% e 0,99993.

Na seção de validação os valores do MSE, MAPE e R^2 foram, respectivamente, 0,000081659, 1,6259% e 0,98493.

5.2 Previsão utilizando Redes Neurais MLP

Foram realizadas previsões para a série Nasdaq usando rede neural MLP, variando a defasagem de 1 até 10 e o número de neurônios na camada oculta de 1 até 25. O método de treinamento foi o Levenberg-Marquardt. A função de ativação da camada oculta foi a sigmóide e da camada de saída a linear. Foram 100 épocas de treinamento.

A figura 5 apresenta o erro quadrático da previsão usando rede neural MLP.

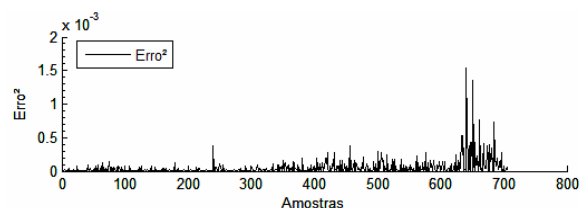


Figura 5 – Erro quadrático da previsão da série Nasdaq usando rede neural MLP para os segmentos de teste e validação

Foi buscado o menor MSE para a seção de teste da série temporal. Desta forma, o melhor valor para a defasagem foi de 3 e o melhor número de neurônios na camada oculta foi de 18.

Tabela 1 – Índices de desempenho para a previsão da série Nasdaq usando a rede neural MLP

Fase	MSE	MAPE	R ²
Estimação	0,000070496	1,2259%	0,99712
Teste	0,000017594	0,67599%	0,99993
Validação	0,00008052	1,5998%	0,98514

Os índices de desempenho encontrados para esses dois parâmetros são apresentados na tabela 1.

5.3 Previsão utilizando Redes Neurais RBF

Foram realizadas previsões para a série Nasdaq usando RN-RBF, variando a defasagem de 1 até 50 e o número de centros de 2 até 50. O *spread* da rede foi fixado em 0,001. A função de base radial usada foi a Gaussiana. O ajuste dos centros foi realizado pelo método *k-means* e o ajuste dos pesos através do algoritmo de mínimos quadrados.

A figura 6 apresenta o erro quadrático da previsão usando rede neural RBF.

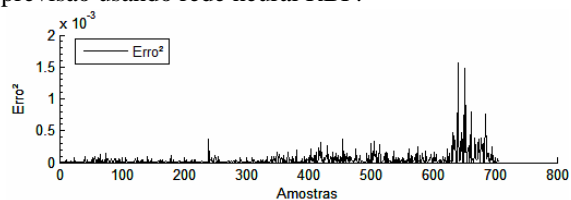


Figura 6 – Erro quadrático da previsão da série Nasdaq usando rede neural RBF para os segmentos de teste e validação

Foi buscado o menor MSE para a seção de teste da série temporal. Desta forma, o melhor valor para a defasagem foi de 1 e o melhor número de centros foi de 2.

Tabela 2 – Índices de desempenho para a previsão da série Nasdaq usando rede neural RBF

Fase	MSE	MAPE	R ²
Estimação	0,000074504	1,2313%	0,99696
Teste	0,000017977	0,68944%	0,99993
Validação	0,000080563	1,6067%	0,98514

Os índices de desempenho encontrados para esses dois parâmetros são apresentados na tabela 2.

6 Conclusão

Neste artigo foram comparados os modelos de previsão Holt-Winters e redes neurais MLP e RBF para a série financeira Nasdaq. A rede neural MLP obteve o melhor desempenho em todos os índices observados (MSE, MAPE e R²) em todos os conjuntos de dados avaliados (teste e validação).

Referências Bibliográficas

- Abelém, A. J. G. (1994). Redes neurais artificiais na previsão de séries temporais. Dissertação de Mestrado, Pós-Graduação em Engenharia Elétrica, PUCRJ, Rio de Janeiro, RJ.
- Castro, M. C. F. (2001). Predição não-linear de séries temporais usando redes neurais RBF por decomposição em componentes principais. Tese de Doutorado, Pós-Graduação em Engenharia Elétrica, Faculdade de Engenharia Elétrica e Computação, UNICAMP, Campinas, SP.
- Corrar, L. J.; Theóphilo, C. R. (2004). Pesquisa operacional para decisão em contabilidade e administração: contabilometria. São Paulo, SP: Atlas.
- Haykin, S. (2001). Redes neurais: princípios e prática. 2. Porto Alegre, RS: Editora Bookman.
- Lewis, C. D. (1997). Demand forecasting and inventory control: a computer aided learning approach. New York: J. Wiley & Sons.
- Makridakis, S. G.; *et al.* (1998). Forecasting: methods and applications. 3rd ed. New York, NY, USA: Wiley.
- Morettin, P. A.; Toloí, C. M. (1987). Séries temporais. 2. Ed., São Paulo: Atual.
- Nasdaq. (2009). NASDAQ composite index methodology. Disponível em: <<http://www.nasdaqtrader.com/content/home/help/indexmethod/CompIndexMethod.pdf>>, acessado em 27/06/2009.
- Paiva, R. P. P. C. (1999). Identificação neuro-difusa: aspectos de interpretabilidade. Dissertação (Mestrado em Engenharia Elétrica), Departamento de Engenharia Informática, Universidade de Coimbra, Coimbra, Portugal.
- Santos, A. A. P. (2005). Previsão não-linear da taxa de câmbio real/dólar utilizando redes neurais e sistemas nebulosos. Dissertação (Mestrado em economia) – Programa de Pós-Graduação em Economia, UFSC, Florianópolis, SC.
- Zhang, G. *et al.* (1998). Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting*, **14**(1): 35-62.