

Redes Neurais para Previsão de Séries Temporais Epidemiológicas

Claudio Bustamante P. de Sá¹ e Flavio F. Nobre²

¹ - CESTE/ENSP, FIOCRUZ. ² - Programa de Engenharia Biomédica, COPPE/UFRJ

Resumo - Propomos o uso de redes neurais artificiais (RNs) como alternativa para prever séries temporais epidemiológicas. Testes com uma série de ocorrência de hepatite B nos E.U.A. obtiveram resultados próximos aos de modelos ARIMA desenvolvidos no CDC.

Abstract - We propose the use of Artificial Neural Networks (ANNs) as an alternative to predict epidemiological time series. Tests with a hepatitis B in the U.S.A. case occurrence series shown results close to those obtained by ARIMA models developed at CDC.

Introdução

Previsões do comportamento de variáveis de saúde são valiosas no contexto de sistemas de Vigilância epidemiológica. Diferentes metodologias, como a de Box-Jenkins e o Amortecimento Exponencial, tem sido propostas para gerar tais predições a partir de séries temporais das variáveis de interesse. Entretanto, estas metodologias, apresentam algumas dificuldades e limitações para seu uso¹.

As Redes Neurais artificiais (RNs) compreendem um conjunto de técnicas matemáticas/computacionais inspiradas nos neurônios. Elas vêm se desenvolvendo rapidamente nos últimos anos e têm sido aplicadas com sucesso em diversas áreas como: simulação e controle de sistemas dinâmicos, processamento de sinais, inteligência artificial e previsão de séries temporais. A maioria das aplicações a séries temporais lidam com variáveis econométricas², sendo também comuns trabalhos de predição de consumo de energia³. Até o presente momento, não encontramos aplicações específicas a séries epidemiológicas.

O objetivo deste trabalho é testar metodologias baseadas em redes neurais como alternativas para predição de séries temporais epidemiológicas.

Metodologia

Os dados utilizados formam a série temporal de incidência de hepatite B em todo território dos E.U.A. entre 1980 a 1989, gentilmente cedida pelo Centers for Disease Control and Prevention (CDC).

Utilizamos redes neurais do tipo "feedforward" com três camadas (entrada, saída e uma camada intermediária), treinadas por backpropagation, onde a saída-alvo é a observação que desejamos prever e as entradas são um grupo de observações passadas.

A definição do grupo de observações que será usado como entrada, não é trivial e foi pouco abordada na literatura até o presente momento. Neste trabalho, propomos uma metodologia onde escolhemos como entradas para previsão de um determinado ponto (x_t) todos os pontos referentes as defasagens (x_{t-n}) cujo valor da autocorrelação seja maior em módulo que os limites de confiança de Bartlett, em nosso caso: x_{t-1} , x_{t-2} , x_{t-3} , x_{t-4} , x_{t-5} , x_{t-6} , x_{t-7} , x_{t-8} , x_{t-9} , x_{t-11} , x_{t-12} , x_{t-13} , x_{t-14} , x_{t-15} , x_{t-16} , x_{t-17} , x_{t-18} , x_{t-19} , x_{t-20} , x_{t-21} , x_{t-31} , x_{t-32} , x_{t-33} , x_{t-34} , x_{t-38} , x_{t-51} , x_{t-52} e x_{t-53} .

Todas as redes implementadas possuíam apenas uma camada intermediária e camadas de saída com apenas um neurônio. Todos os neurônios tinham função de ativação logística. As séries foram normalizadas no intervalo $[0,1;0,9]$ e as sinapses inicializadas com valores aleatório pequenos. Os pares de treinamento (saídas e seus respectivos vetores de entradas) foram apresentados a rede de três modos distintos: aleatoriamente, sequencialmente e em lote. Em todos os casos usamos coeficientes de aprendizado (α) adaptativos com correção de $[1-(1/\gamma)]$. Nos treinamentos aleatório e seqüencial incluímos ainda um termo de momento (β). Para o treinamento em lote escolhemos $\alpha=0,02$ (inicial) e $\gamma=40$. Nas outras modalidades usamos $\alpha=0,1$ (inicial), $\gamma=5000$ e $\beta=0,9$.

Os dados foram divididos em três grupos: conjunto de treinamento (anos 1980-1987), conjunto de teste (1988) e conjunto avaliação das previsões (1989). As redes foram treinadas por 50000 épocas (400 para o treinamento em lote). O treinamento foi implementado por uma série de rotinas escritas para o pacote de processamento numérico MATLABTM (The MathWorks Inc.).

Resultados

Para cada rede treinada, foram feitas previsões para cinco horizontes diferentes: 1 semana, 4 semanas 13 semanas 26 semanas e 52

semanas. Utilizamos como figura de mérito para as previsões o erro médio absoluto percentual (EMAP):

$$\text{EMAP} = \frac{\sum_{i=1}^n |x_i - \hat{x}_i|}{n} \times 100$$

onde x_i são as observações do conjunto separado para avaliar as previsões e \hat{x}_i as previsões correspondentes.

Para comparação apresentamos também resultados obtidos por modelos ARIMA desenvolvido no CDC para as mesmas séries⁴ (figura 1).

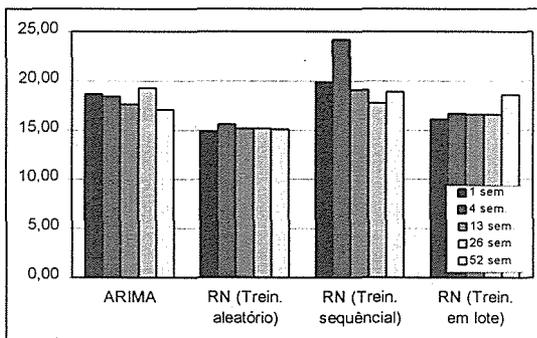


Figura 1: Erros de previsão (EMAP) na série de hepatite B.

Discussão

Utilizamos em nosso trabalho apenas redes do tipo “feedforward”, treinadas por backpropagation. Alguns autores⁵ sugerem que o uso de redes neurais realimentadas é mais adequado para previsão. Entretanto, os demais trabalhos revistos utilizaram redes “feedforward”, que são mais bem conhecidas e de mais fácil implementação.

Devido ao tempo computacional consumido pelo processo de treinamento, fixamos um tamanho para as camadas intermediárias ao invés de testar diversas possibilidades. Optamos por usar o mesmo número de unidades empregado na camada de entrada, com a idéia de garantir que não houvesse “underfitting”. Para evitar o “overfitting”, selecionamos, para cada rede, o conjunto de pesos que obteve o menor erro no conjunto de teste ao final treinamento.

O treinamento com apresentação aleatória obteve os melhores resultados, superiores aos do modelo ARIMA. O treinamento em lote teve um desempenho um pouco inferior, próximo aos do modelo ARIMA e o treinamento seqüencial obteve os piores resultados.

Referências

- ¹ Nobre, F. F. e Stroup, D. F. (1994), “A Monitoring System to Detect Changes in Public Health Surveillance Data”, *Int. Jour. of Epidemiology*, v. 23, 2, 408-418.
- ² Tang, Z. *et al.* (1991), “Time series forecasting using neural networks vs. Box-Jenkins methodology”, *Simulation*, vol. 57, no. 5, 303-310.
- ³ Peng, T. *et al.* (1992), “Advancement in the Application of Neural Networks for Short-Term Load Forecasting”, *IEEE Trans. on Power Systems*, vol. 7, no. 1.
- ⁴ Williamson, G. *et al.* (1991), “A Monitoring System for Improved Public Health Surveillance” apresentado no *World Fed. Pub. Health Assoc. 6th Int. Congress*.
- ⁵ Connor, J. *et al.* (1994), “Recurrent Neural Networks and Robust Time Series Prediction”, *IEEE Transactions on Neural Networks.*, vol. 2, no. 2, 240-254.