

PREVISÃO DAS CHUVAS DO MÊS DE OUTUBRO EM PELOTAS, RS, USANDO UMA REDE NEURAL MLP COM CONFIGURAÇÃO OTIMIZADA POR ALGORITMOS GENÉTICOS.

Márcio Porto Basgalupp¹, Mário Lúcio Mesquita Machado², João Artur de Souza³, João Baptista da Silva⁴

ABSTRACT - The behavior of the meteorological elements, such as temperature, rainfall, wind, etc, has been, along time, object of study by man. The importance of these elements, as factors of human survival, had arisen in the scientists of the whole world a large interest in its interpretation, simulation and forecasting. Several methodologies such as spectral analysis, Box & Jenkins models, Wavelets, Neural Network, etc. have been used in modeling that climatic elements. This paper is about the use of Genetic Algorithms (AG) in order to enhance the definition of the Multilayer Perceptron Neural Network architecture, within a practical application of time series analysis. The time series are that from October monthly precipitation, at Pelotas, RS. The AG optimization is intended to support a better capacity in generalization for the forecasts, minimizing errors. The results of this work emphasize the efficiency of the technique used (MLP with GA) in the rainfall forecasting of October month at Pelotas, RS.

INTRODUÇÃO

O estudo climatológico das diversas variáveis do tempo é de extrema importância, tendo em vista o impacto ambiental que a anomalia destas componentes provocam no clima regional. Análises dos comportamentos das séries climáticas desses elementos, visando destacar possíveis periodicidades existentes, são fundamentais para o planejamento de inúmeras atividades agrícolas, econômicas, sociais, dentre outras.

Existem várias metodologias para a modelagem destes elementos meteorológicos, entre elas: Análise Harmônica, Análise Espectral, Modelos de Box & Jenkins, *Wavelets*, etc. Baptista da Silva et al. (2004) utilizaram a técnica da análise harmônica para a modelagem das chuvas mensais em Pelotas, RS, obtendo resultados satisfatórios.

Uma rede neural artificial é, de maneira geral, um sistema construído para modelar a forma pela qual o cérebro realiza uma determinada tarefa ou função. Assim, o conhecimento para a solução de um determinado problema é obtido através de um processo de aprendizagem (Haykin, 2001).

Algoritmos Genéticos (AG's) são estratégias de busca desenvolvidas por Holland (1975), que baseou esta técnica no mecanismo de seleção natural proposto por Darwin, no qual o mais apto tem o direito de se reproduzir e sobreviver (Goldberg, 1989).

As redes neurais artificiais também podem ser utilizadas como técnica de modelagem de dados. Entretanto, possui sua arquitetura vinculada ao tipo de problema a que se destina. Assim, é importantíssimo descobrir a arquitetura que forneça a solução mais eficiente de um problema específico. Uma técnica que vem sendo utilizada para encontrar essa arquitetura ótima e tem gerado bons resultados nesse sentido é a de Algoritmos Genéticos.

Este trabalho tem como objetivo avaliar as técnicas de redes neurais artificiais, com sua

configuração otimizada por algoritmos genéticos, na modelagem das chuvas do mês de outubro em Pelotas, RS.

MATERIAL E MÉTODOS

Foram utilizados dados de umidade relativa, temperatura média, velocidade do vento e de chuvas mensais no município de Pelotas, RS, referentes aos anos de 1960 a 1999, sendo que os 8 (oito) últimos anos (1992/99) foram utilizados para avaliar a previsão, não participando da construção do modelo. Esses dados foram coletados na Estação Agroclimatológica de Pelotas, com latitude 31°52'S, longitude 52°21'24" W.GRW e altitude 13,24m.

Para ensinar a rede neural a calcular o valor da chuva no mês de outubro para um ano $x+1$, foram fornecidos à entrada da rede os valores das variáveis: velocidade do vento, chuva, temperatura média e umidade relativa dos meses de julho, agosto, setembro e outubro do ano x , totalizando 16 (dezesesseis) valores. Assim, concluímos que a camada de entrada da rede neural possui 16 (dezesesseis) neurônios.

Um AG trabalha iterativamente sobre um grupo de soluções candidatas para o problema, denominado população. A cada iteração realizada são geradas novas soluções a partir de determinadas operações, que serão realizadas sobre os membros da população atual. Cada membro da população é denominado cromossomo, recebendo esse nome pelo fato de codificar os parâmetros da solução do problema em uma string de símbolos de um determinado alfabeto, similarmente a uma seqüência de DNA. Usualmente os símbolos do alfabeto são 0's e 1's. Cada iteração de um AG é chamada de geração.

A topologia de rede neural empregada para fazer a modelagem em questão é a MLP (*Multilayer Perceptron*) com uma camada oculta (Figura 1). Considerando-se que em uma rede MLP as unidades de uma camada são totalmente conectadas às unidades da camada seguinte, pode-se notar que não existe a necessidade de representar detalhes de conectividade da rede nos cromossomos. Os cromossomos possuem a seguinte seqüência:

NNCO	FACO	FACS	TA
------	------	------	----

NNCO = n° de neurônios na camada oculta (4-30);

FACO = função de ativação da camada oculta (1-3);

FACS = função de ativação da camada de saída (1-3);

TA = taxa de aprendizagem (0-1).

Para a camada oculta foram definidas 3 possíveis funções de ativação: linear pura, tangente hiperbólica e sigmoideal, numeradas nessa ordem.

Considerando que o objetivo da rede neural, neste trabalho, é realizar a previsão da chuva para o mês de outubro do ano seguinte, o teste foi realizado de maneira contínua. Assim, após a rede prever o valor da chuva de outubro do próximo ano, o valor observado nesse ano é incluído ao conjunto de treinamento e a rede é treinada um pouco mais para, após, prever o

¹ Mestrando em Ciência da Computação pela PUC-RS (basgalupp@inf.pucrs.br)

² Graduando em Ciência da Computação pela UFPel (krazny@ufpel.edu.br)

³ Doutor em Engenharia de Produção pela UFSC (jartur@gmail.com)

⁴ Livre Docente, Doutor em Ciências pela UFPel (jsilva@ufpel.edu.br)

valor do próximo ano. Esse processo se repete até a previsão do último valor do conjunto de testes.

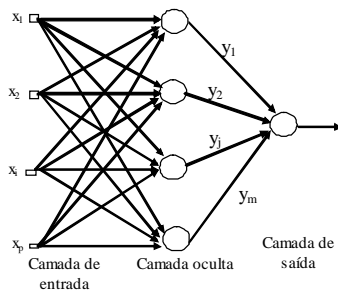


Figura 1. MLP com uma camada oculta e um neurônio na camada de saída.

Os operadores genéticos empregados foram o *crossover* (cujo objetivo é cruzar indivíduos) de um ponto e a mutação (o valor de um gene escolhido para ser alterado é trocado por um valor aleatório dentro de uma faixa considerada). O operador de *crossover* é empregado sobre uma lista de indivíduos (membros da população) que se mostraram mais apta numa tentativa de usar um elitismo. O tamanho escolhido para a população foi de 20 indivíduos e a nova população gerada a cada iteração substitui completamente a antiga. Esse processo de geração da nova população foi repetido 50 vezes. O melhor elemento de cada geração é salvo, proporcionando uma comparação. O que indica se um indivíduo é melhor que o outro é o valor de sua função de *fitness* (cujas tarefas é atribuir um valor numérico a cada cromossomo, indicando o “quão” apto é o indivíduo) para os dados de teste.

A função de *fitness* empregada foi o EQRM (Erro Quadrático Relativo Médio):

$$EQRM = \left[\frac{1}{N} \sum_{i=1}^N \left(\frac{C_{obsi} - C_{RNi}}{C_{obsi}} \right)^2 \right]^{1/2}$$

onde C_{obs} é a chuva observada, C_{RN} é a chuva calculada pela rede neural e N é o número de observações.

Como o objetivo é encontrar o menor EQRM possível, este problema se caracteriza como de minimização.

Para avaliar estatisticamente o modelo como predictor, foi calculado o valor do teste t , onde

$$t = \sqrt{\frac{(n-1)MBE^2}{RMSE^2 - MBE^2}}, \text{ sendo } MBE = \frac{1}{n} \sum_{i=1}^n (C_{obsi} - C_{rni}) \text{ e}$$

$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (C_{obsi} - C_{rni})^2}$, entre os valores estimados pelo modelo e os valores observados (Togrul & Togrul, 2002).

RESULTADOS E DISCUSSÕES

Todo o processo durou aproximadamente 70 horas e a configuração do indivíduo “vencedor” foi a seguinte:

NNCO = 9	FACO = 2	FACS = 1	TA = 0,09
----------	----------	----------	-----------

O resultado da rede neural gerada por essa configuração para os dados dos oito anos utilizados para a previsão, de 1992 a 1999, é apresentado na Figura 2.

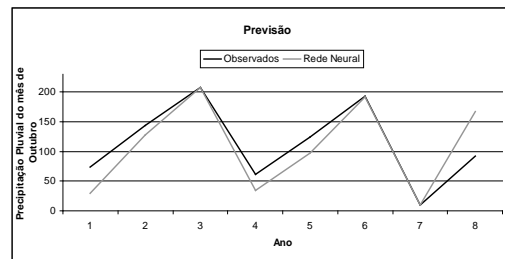


Figura 2. Valores observados e estimados pelo modelo de redes neurais, para os oito anos de previsão.

Observa-se que o modelo de rede neural acompanhou a tendência dos dados observados e, dos 8 (oito) valores testados, 3 (três) foram estimados com exatidão.

O valor do EQRM (função de *fitness*), foi igual a 0,4016, o que é considerado um valor baixo, exprimindo a boa aproximação do modelo de redes neurais aos dados observados.

O valor de t calculado, $t = 0,4156$, permite concluir que o modelo é um bom predictor da variável em estudo, pois esse resultado é não-significativo ao nível de 5% ($t_{0,05; 7} = 2,36$).

CONCLUSÕES

Os algoritmos genéticos se aplicaram bem à otimização da configuração da rede neural, podendo-se adicionar outros parâmetros nos indivíduos (número de épocas e algoritmos de treinamento).

Ficou evidenciada a capacidade das redes neurais em prever os valores das chuvas do mês de outubro em Pelotas, RS. Pode-se inferir que esta técnica também produzirá bons resultados para os outros meses do ano.

REFERÊNCIAS

- Baptista da Silva, J.; Basgalupp, M. P.; Souza, D. O. et al. Modelagem e previsão das precipitações pluviométricas mensais em Pelotas, RS: Análise harmônica. In: X Reunion Argentina y IV Latinoamericana de Agrometeorología, Mar del Plata - Argentina. Agrometeorología y seguridad alimentaria en América Latina, 2004.
- Goldberg, D. E. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, p. 412, 1989.
- Haykin, S. Redes Neurais: princípios e prática. Bookman, 2ª edição, 2001. 900p.
- Holland, J. H. Adaptation in Natural and Artificial Systems, University of Michigan Press: Ann Arbor, MI, 1975.
- Lucas, D. C. Algoritmos Genéticos: um estudo de seus conceitos fundamentais e aplicação ao problema da grade horária. Trabalho de conclusão do curso de Bacharelado em Ciência da Computação. Universidade Federal de Pelotas. Pelotas - RS, 2000.
- Togrul, I.T., Togrul, H. Global solar radiation over Turkey: comparison of predicted and measured data. Renewable Energy, v.25, p.55-67. 2000.