

# **Estimativa De Padrões Espaço-Temporais Utilizando Redes Neurais Artificiais do Tipo *Multilayer Perceptron* com Aplicação em Física Solar**

**Maria Conceição de Andrade <sup>\*1</sup>, Atair RiosNeto <sup>\*\*2</sup>, Reinaldo R. Rosa <sup>\*\*\*3</sup>**

**(1)Área de Inteligência Artificial**

**Divisão de Astrofísica**

**(2)Área de Inteligência Artificial**

**(3)Área de Computação Científica**

**Laboratório Associado de Computação e Matemática Aplicada**

**Instituto Nacional de Pesquisas Espaciais (INPE)**

**(\*)Mestrado, e-mail: [con@das.inpe.br](mailto:con@das.inpe.br) (\*\*) Orientadores**

## **Resumo**

Este trabalho se propõe a avaliar e comparar o desempenho de redes neurais artificiais utilizando os algoritmos Backpropagation, Levenberg-Marquardt e Filtro de Kalman para o treinamento de redes neurais do tipo Multilayer Perceptron aplicadas a problema de estimativa de padrões espaço-temporais. O estudo de caso consiste na aplicação do Multilayer Perceptron para caracterização da evolução de regiões ativas solares observadas pelo satélite YOHKOH. O teste de desempenho da aplicação será principalmente realizado através da técnica de Análise de Padrões Gradientes (APG) utilizando o operador para caracterização de Fragmentação Assimétrica da Amplitude (FAA). o FAA será utilizado, pela primeira vez, para o teste de desempenho de redes neurais aplicadas a estimação de padrões espaço-temporais.

**Palavras-Chave:** *redes neurais artificiais, treinamento supervisionado, estimação de séries espaço-temporais, análise de padrões gradientes, física solar.*

## **1. Introdução**

Investigamos a implementação e análise de desempenho de Redes Neurais Artificiais (RNA) na estimação de padrões espaço-temporais da coroa solar. Os padrões são visualizados a partir de imagens de satélite da dinâmica espaço-temporal das regiões ativas solares [7]. Para estimativa desses padrões serão utilizados os algoritmos de treinamento Backpropagation, Levenberg-Marquardt[1] e Filtragem de Kalman[4], bem como serão analisados os resultados de cada um deles. Com base nestas imagens, estudos de previsão de eventos solares em grande escala poderão ser feitos permitindo assim prever possíveis efeitos posteriores nas regiões mais próximas da Terra (tempestades magnéticas e irregularidades ionosféricas). A técnica poderá ser utilizada como ferramenta fenomenológica para previsão da atividade solar dentro do Programa Latino-Americano para Investigação do Clima Espacial (PLAICE) [6] e para o Sistema de Análise Fenomenológica (SAF), que está sendo desenvolvido pelo Núcleo de Simulação e Análise de Sistemas Complexos (NUSASC) [8]. As respostas da rede são avaliadas usando o operador computacional de matrizes FAA (*Fragmentação Assimétrica da Amplitude*) [5]. Aplica-se o operador em ambas imagens, ou seja, na resposta da rede neural (estimativa) e na imagem real. A expectativa de obter bons resultados com RNA é baseada no fato de que nos últimos anos elas têm mostrado ser uma efetiva ferramenta em previsões no comportamento de séries temporais, e têm consequentemente sido aplicadas na previsão de várias séries temporais solar-terrestres, como pode ser visto em [3]. Neste trabalho estaremos estendendo esta aplicação para o domínio espaço-temporal (análise de sequência de imagens). A utilização de RNA em previsão de imagens é atrativa por ter a capacidade de "aprender" padrões que por sua complexidade, tornam-se difíceis de ser definidos por outras técnicas, com a habilidade de integrar informações a partir de amostras e incorporar novas características, sem degradar o conhecimento adquirido anteriormente. O *Multilayer Perceptron* tem sido aplicado com sucesso para resolver problemas diversos através de treinamentos supervisionados com o popular algoritmo *Backpropagation*.

## **2. Algoritmos de Treinamento**

O algoritmo Backpropagation é o mais difundido e amplamente usado, a implementação de uma rede baseada nesse algoritmo pode demandar muitos passos e resultar num tempo de treinamento muito longo.

Frente a esta questão, existe um número de variações no algoritmo Backpropagation básico, que são baseadas em outras técnicas de otimização, tais como, gradiente conjugado, método de Newton, Filtragem de Kalman. Neste trabalho foram escolhidos para treinamento da rede os algoritmos, Backpropagation, Levenberg-Marquardt[1] e Filtragem de Kalman[4]. Será apresentada a performance de cada um deles para o problema proposto, tendo em vista os resultados alcançados.

### 2.1 Algoritmo Backpropagation

O Backpropagation é uma técnica específica para uma rede feedforward multi-camadas que implementa o gradiente descendente no espaço dos pesos. Quando treinada adequadamente, a rede Backpropagation tende a dar respostas razoáveis diante de entradas ainda não conhecidas.

### 2.2 Algoritmo Levenberg-Marquardt

Levenberg Marquardt [1] é uma aproximação ao método de Newton, uma técnica de otimização mais potente do que o gradiente descendente. A regra de atualização dos pesos de Levenberg-Marquardt é:

$$\Delta w = (J^T J + \mu I)^{-1} J^T e$$

onde  $J$  é a matriz Jacobiana das derivadas de cada erro para cada peso,  $\mu$  é um escalar, e  $e$  é o vetor dos erros. Se o escalar  $\mu$  for muito grande, a expressão acima aproxima-se do gradiente descendente, caso contrário a expressão torna-se o método de Gauss-Newton. O método de Gauss-Newton é mais rápido e mais preciso, próximo de um erro mínimo, portanto a meta é transferir para o método de Gauss-Newton. Então  $\mu$  será acrescido ou decrescido depois de cada passo conforme for o caso [9].

### 2.3 Algoritmo Filtro de Kalman Estendido

O Filtro de Kalman é um algoritmo usado para estimar estados de um sistema dinâmico a partir de medidas ruidosas; torna-se possível, portanto, o seu uso em redes neurais, uma vez que ela busca ajustar (estimar) pesos que encontrem um modelo computacional para um dado mapeamento de conjunto de dados do tipo:

$$\{(x(t), y(t)) : y(t) = f(x(t), t = 1, 2, \dots, m)\}$$

Os pesos das redes neurais são os estados a serem estimados e as saídas da rede são as medidas a partir das quais o Filtro de Kalman utiliza para estimação:

$$\hat{y}(t) = \hat{f}(x(t), w)$$

onde:  $w$  é o vetor de pesos sinápticos do neurônio,  $x(t)$  são os dados de entrada da rede.

Na solução proposta [4], é feita uma perturbação linear em uma iteração  $i$ , transformando o treinamento no seguinte problema de estimação:

$$\begin{aligned} \varpi &= w(i) + \bar{e} \\ \alpha(i)[y(t) - \bar{y}(t, i)] &\cong \hat{f}_w(x(t), \bar{w}(i))[w(i) - \bar{w}(i)] + v(t) \end{aligned}$$

onde:  $i = 1, 2, \dots, I$ ;  $\varpi$  = é a estimativa a priori de  $w$ , vindo da iteração anterior começando com  $\varpi(1) = \varpi$ ;  $\bar{y}(t, i) = \hat{f}(x(t), \varpi(i))$ ;  $\hat{f}_w(x(t), \varpi(i))$  é a matriz das derivadas parciais com respeito a  $w$ ;  $0 < \alpha(i) \leq 1$  é um parâmetro a ser ajustado para garantir a hipótese de perturbação linear;  $\bar{e}$  a distribuição de erro a priori, e  $v(t)$  o erro na aproximação das saídas da rede, são variáveis aleatórias gaussianas.

## 3. Operador de Fragmentação Assimétrica de Amplitudes

O operador FAA é capaz de quantificar localmente as quebras de simetria no campo gradiente de uma imagem. Transforma uma matriz  $M$  em um campo de triangulação com  $L$  pontos e  $I$  linhas, onde os  $L$  pontos correspondem ao número de vetores assimétricos do campo gradiente da matriz e  $I$  linhas correspondem à quantidade de linhas da triangulação entre os  $L$  pontos. A medida do grau de fragmentação assimétrica é dada pelo valor  $F_A = (I-L)/L$ . Este parâmetro apresenta alta sensibilidade para caracterizar pequenas alterações no campo gradiente de uma matriz, refletindo as possíveis quebras de simetria que possam ocorrer no domínio espaço-temporal[5].

## 4. Conjunto de Padrões e Metodologia de Treinamento

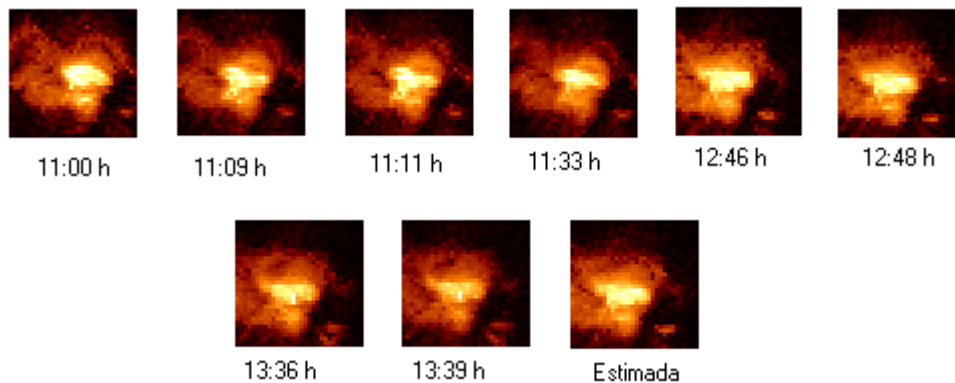
Os padrões usados neste trabalho, são imagens da coroa solar em Raios-X moles obtidas pelo satélite YOHKOH[10]. Foi usada uma série de imagens com pouca alteração nos padrões. Estas imagens são disponibilizadas em formato ASCII (American Standard Code for Information Interchange), no tamanho de 128x128 pixels. Para o treinamento as imagens foram reamostradas em 32x32 pixels e normalizadas. O conjunto de dados de treinamento é formado por 5 padrões, cada um composto de uma entrada e uma

saída. A imagem de saída do primeiro padrão é a entrada do segundo padrão, e assim sucessivamente, até o último. Cada pixel de um padrão de entrada é considerado um neurônio, totalizando 1024 neurônios na entrada, e 1024 neurônios na saída, totalmente conectados, funções de ativação, sigmóide e linear nas camadas escondida e saída respectivamente.

## 5. Resultados alcançados

Algoritmo Backpropagation com a técnica de *momentum*: treinamento feito com a arquitetura e padrões acima descritos. O erro do treinamento alcançado foi de  $8.4 \times 10^{-4}$ . A figura 1 mostra a série temporal dos padrões apresentados no treinamento. A figura 2 apresenta a imagem usada para teste, (posterior a série temporal), a imagem que a rede neural deve estimar e a resposta da rede neural.

**Figura 1 – Padrões de treinamento**



**Figura 2 - Padrão de teste; imagem a ser estimada e imagem estimada**

O algoritmo Levenber-Marquardt não apresentou bom desempenho para este tipo de problema. O algoritmo utiliza muita memória no cálculo de correção do peso, mesmo em um pentium-III 600MHz, com 380 Mb de RAM.

O algoritmo Filtro de Kalman continua em fase de treinamento e até o momento não obtivemos resposta da rede.

## 6. Conclusões

A partir dos resultados preliminares concluímos que a metodologia é promissora para este fim, desde que seja bem calibrada para o problema em questão. O algoritmo Levenberg\_Marquardt apesar de ser um algoritmo robusto, para este problema não é eficaz. Até o momento o algoritmo Backpropagation foi o que apresentou melhor performance para este estudo de caso.

## Referências

- [1] Hagan M. T., Menhaj M. B., "Training Feedforward Networks with the Marquardt Algorithm", IEEE Transactions on Neural Networks, Vol. 5, No. 6, November 1994
- [2] Haykin, S. "Neural Networks – A Comprehensive Foundation", 1994 Prentice-Hall, Inc.
- [3] Macpherson K. P., Conway A. J., Brown J. C. "Prediction of solar and geomagnetic activity data using neural networks", Journal of Geophysical Research, vol. 100, NO. A11, Pages 21735-21744, Novembro 1995
- [4] Rios Neto, A., "Kalman Filtering Stochastic Optimal Estimation Algorithm and Usual Backpropagation in Neural Nets Training", Proc. of Second Brazilian Cong. in Neural Net., Curitiba, Paraná, Brasil, 139-144, 1995
- [5] Rosa, R. R., Sharma, A.S., Valdivia, J.A., "Characterization of asymmetric fragmentation patterns in spatially extended systems", International Journal of Modern Physics C, Vol. 10, No 1 (1999) 147-163
- [6] Rosa, R.R. et al, "Computational operators for dynamical complex pattern recognition", in EPS- Modelling Collective Phenomena in Complex Systems, V. 22F, Pub. by European Physical Society, pp. 304-305, 1998
- [7] Rosa, R. R.; Sawant H.S.; Saito J. H., "Tomografia Espectral Solar utilizando DSPs em Arquitetura Paralela", Instituto Nacional de Pesquisas Espaciais – Divisão de Geofísica Espacial e Astrofísica
- [8] Rosa, R. R.; Rodrigues, C.R., "Caracterização de Regimes Complexos em sistemas Dissipativos Auto-organizados", Projeto Fapesp Proc. 97/13374-1, 1997
- [9] The MathWorks, Inc., Neural Network Toolbox – User's Guide – 1994
- [10] <http://www.solar.isas.ac.jp>