Análise de um aproximador funcional utilizando as Redes Neurais artificiais MLP treinada com o algoritmo Backpropagation

Martinez, S. C. [1], Ferrara, L. F. P. [2], Mario, M. C. [2].

[1]Aluna do Curso de Mestrado na Universidade Santa Cecília, Santos, BR.
[2] Professor do Curso de Mestrado na Universidade Santa Cecília, Santos, BR.

As redes neurais artificiais são modelos matemáticos que possuem características análogas às redes neurais biológicas. O funcionamento é baseado na estrutura do neurônio real onde cada neurônio recebe informações (sinais elétricos) dos seus inúmeros dendritos, e estes sinais podem ser atenuados ou amplificados, dependendo do dendrito a que está associado. Cada dendrito possui um peso que armazena todo o conhecimento de uma rede neural e a sua atualização é o que se entende por aprendizado. O objetivo desse trabalho é realizar um estudo sobre as características das redes neurais e do algoritmo *backpropagation* para aproximar funções contínuas com exatidão. Foi desenvolvido um simulador que realiza o treinamento e o acompanhamento gráfico. A rede neural utilizada apresentou resultados satisfatórios durante o treinamento, demonstrando uma excelente capacidade de generalização. Corroborando com a ideia de modelagem de funções sem a necessidade do conhecimento prévio da mesma, a aplicabilidade desse sistema pode se estender a área de controle.

Palavras chave: rede neural, aproximador funcional, retro propagação, modelos matemáticos, generalização.

Analysis of a function approximator using Artificial Neural Networks MLP trained with the Backpropagation algorithm.

The artificial neural networks are mathematical models which have analogous features like the biological neural networks. The operation is based on the structure of the real neuron where each one receives information (electrical signals) from the innumerous dendrites, and these signals can be attenuated or amplified, depending on the dendrite to which it is associated. Each dendrite has a weight which stores all the knowledge of a neural network and its updating is what one understands by learning process. The goal of this work is to perform a study on the characteristics of the neural networks and the backpropagation algorithm for approaching continuous functions with accuracy. It was developed a simulator that performs the training and follow-up of the graph. The neural network used showed satisfactory results during training, demonstrating an excellent capacity of generalization. Corroborating with the idea of modeling of functions without of prior knowledge of the same, the applicability of this system can be extended to the control area.

Key words: *neural network, function approximator, backpropagation, mathematical models, generalization.*

INTRODUÇÃO

Os estudos sobre inteligência artificial se tonaram um vasto campo de pesquisa, estes tipos de sistemas computacionais inteligentes são capazes de resolver problemas simples e complexos, baseado no conceito da inteligência do ser humano. Diversos tipos de sistemas "inteligentes" podem ser criados, tais como, sistemas especialistas, sistemas baseados em redes neurais, em algoritmos genéticos, na lógica *fuzzy*, entre outros. (CAMPOS, 2004 & TONSIG, 2009)

Os primeiros estudos sobre redes neurais vieram em 1943 através do trabalho de Pitts e McCulloch, eles propuseram a utilização das redes de neurônios sendo expressas como uma unidade de processamento binária e provaram que estas unidades executavam diversas operações lógicas. Desta forma eles definiram o "neurônio formal" copiando a ideia funcional da rede neural biológica, e baseados nos resultados da neurobiologia propuseram arquiteturas para a realização das funções lógicas. (CAMPOS, 2004) O modelo do neurônio artificial proposto por Pitts e McCulloch apresenta apenas uma saída, ou seja, a função da soma (*threshold*) do valor de suas diversas entradas. (TONSIG, 2009)

As redes neurais artificiais são sistemas compostos por unidades de processamento, chamadas de neurônios, os quais apresentam características análogas a dos neurônios reais, ou seja, possuem um algoritmo matemático que procura simular a forma, o comportamento e as funções de um neurônio biológico. De uma maneira simplificada podemos dizer que cada neurônio recebe informações (sinais elétricos), dos seus inúmeros dendritos, e estes sinais podem ser atenuados ou amplificados, pois cada dendrito está associado a um peso sináptico. Esses pesos armazenam todo o conhecimento de uma rede neural e a sua atualização é o que se entende por aprendizado. (MACHADO, 2012 & CAMPOS, 2004)

A arquitetura de uma rede neural pode possuir camadas intermediárias ou não. Com as camadas intermediárias o poder computacional aumenta expressivamente, pois as redes multicamadas podem tratar dados que não são linearmente separáveis. A rede de neurônios multicamadas usualmente é classificada em três grupos: as camadas de entradas, de saídas e intermediárias ou ocultas. (NASCIMENTO, 2004)

O *backpropagation* é um algoritmo para treinamento de redes multicamadas, tendo como base o aprendizado supervisionado por correção de erros. Esse algoritmo segue uma sequencia de dois passos, primeiro um padrão é apresentado à entrada e sua resposta é propagada como entrada para as camadas seguintes, camada por camada, até que a resposta seja produzida pela camada de saída. No segundo passo a saída obtida é comparada a saída desejada, caso o valor não seja o esperado o erro é calculado e propagado até a camada de entrada, os pesos sinápticos das camadas internas vão sendo modificados conforme o erro é retro propagado. (FERNANDES, 2004)

A rede neural artificial é uma ferramenta atraente para a resolução de problemas, como reconhecimento de padrões, função de aproximação, mapeamento do comportamento de sistemas não lineares, entre outros. (SILVA, 2011)

O intuito dessa pesquisa é a realização do estudo das redes neurais artificiais como ferramenta para a aproximação de funções; utilizando uma rede neural artificial multicamada treinada através do algoritmo *backpropagation*.

MATERIAIS E MÉTODOS

Para o problema proposto foi utilizada uma Rede Neural montada com três camadas: a camada de entrada composta de um neurônio, a camada de saída composta de um neurônio e a camada intermediária composta de seis neurônios. A arquitetura da rede neural utilizada pode ser visualizada na figura1.

Na simulação deste estudo foi adotada a função trigonométrica $Y = sen(2x)sen(x)$. Como a amplitude da função varia de –0,8 até +0,8 utilizaremos a sigmoide bipolar (-1, 1) como função de ativação nos neurônios contidos na camada escondida e na saída.

Função Sigmoide Bipolar: $f = \frac{2}{1+e^{-x}} - 1$.

O método escolhido para atualização dos pesos foi o *backpropagation* (Regra Delta Generalizada). O Algoritmo de treinamento e o acompanhamento gráfico do treinamento foram implementados utilizando Visual Basic.

O treinamento da rede foi reduzido a apenas um período $T = 2\pi\ rad$. Para que a Rede responda satisfatoriamente a valores superiores a 2π seria necessário um treinamento específico ou a adoção de algum artificio que reduza a entrada da rede ao valor significativo do primeiro período.

No processo de treinamento da rede serão inseridos, em sua entrada, um conjunto de 61 padrões diferentes compreendidos no intervalo de 0 até 2π em intervalos de $\frac{\pi}{30}$.

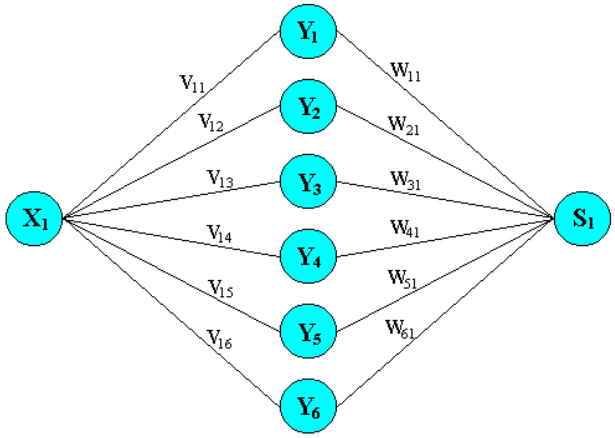

Figura 1 - Arquitetura da rede neural.

RESULTADOS E DISCUSSÃO

Foi possível observar em um primeiro momento que em 10000 ciclos de treinamento, com a taxa de aprendizado em 0.1 a taxa de variação foi muito pequena, assim para um melhor acompanhamento do processo de aprendizagem foi adotado treinamento supervisionado juntamente com a estratégia de exibir os resultados dos pesos e o desempenho da rede, dentro de um período de ciclos pré-estabelecidos, e possibilitar a alteração da taxa de aprendizado caso se

julgue necessário. Os pesos do último ciclo de treinamento serão mantidos. Inicialmente foi adotada a taxa de aprendizado de 0.3 para períodos de 2000 ciclos.

Com o aumento da taxa de aprendizado ocorreu uma aceleração no processo. Pode-se observar na figura 2 que após um período de mais 14000 ciclos (curva verde) a rede já apresentou uma resposta gráfica semelhante à representada pelos padrões de entrada. Nos períodos seguintes a rede aproxima-se cada vez mais dos padrões fornecidos. Nessa etapa será mantida a taxa de aprendizado em 0.3 e o período de treinamento será aumentado para períodos de 5000 ciclos.

Após um período de treinamento de 50000 ciclos é possível observar na figura 3 que a rede está com respostas "próximas" do padrão apresentado. Com o intuito de aumentar a precisão e acelerar o processo, já que nos últimos 30000 ciclos foram obtidos poucas variações, o período de treinamento foi aumentado para períodos de 50000 ciclos e a taxa de aprendizado para 1.5, o resultado pode ser observado na figura 4.

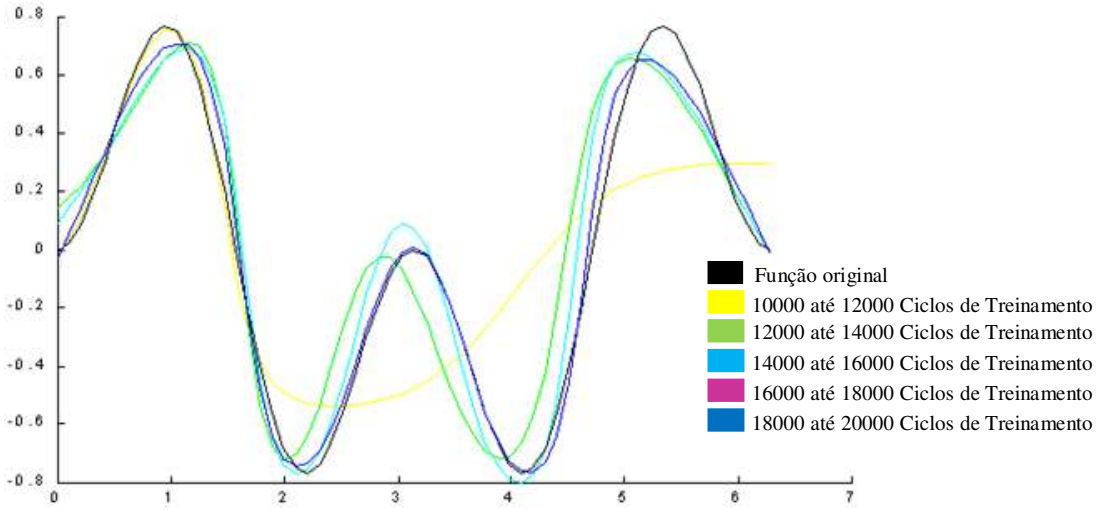

Figura 2 - Sobreposição das figuras geradas no treinamento de 10000 a 20000, com a taxa de aprendizado em 0.3.

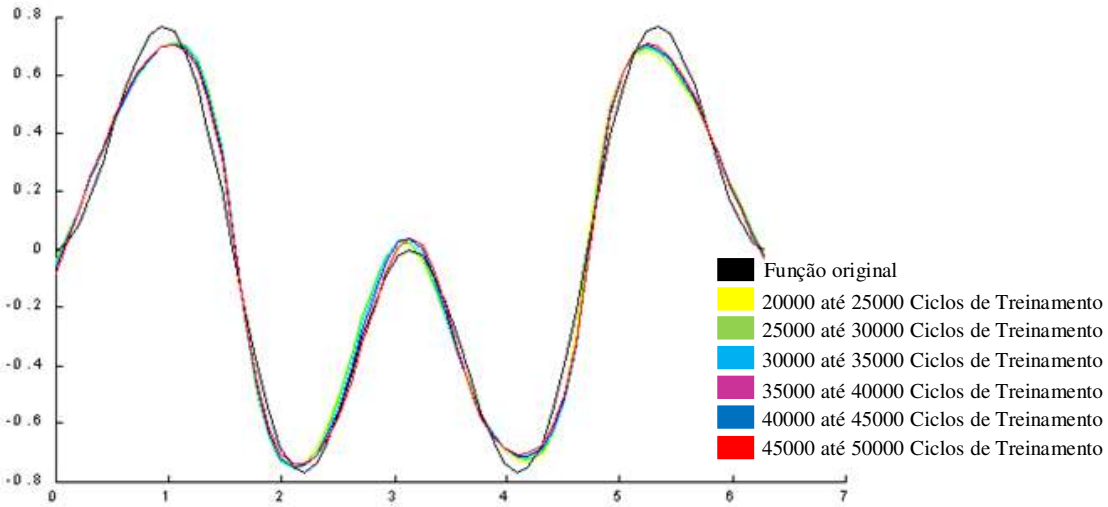

Figura 3 - Sobreposição das figuras geradas no treinamento de 20000 a 50000, com a taxa de aprendizado em 0.3.

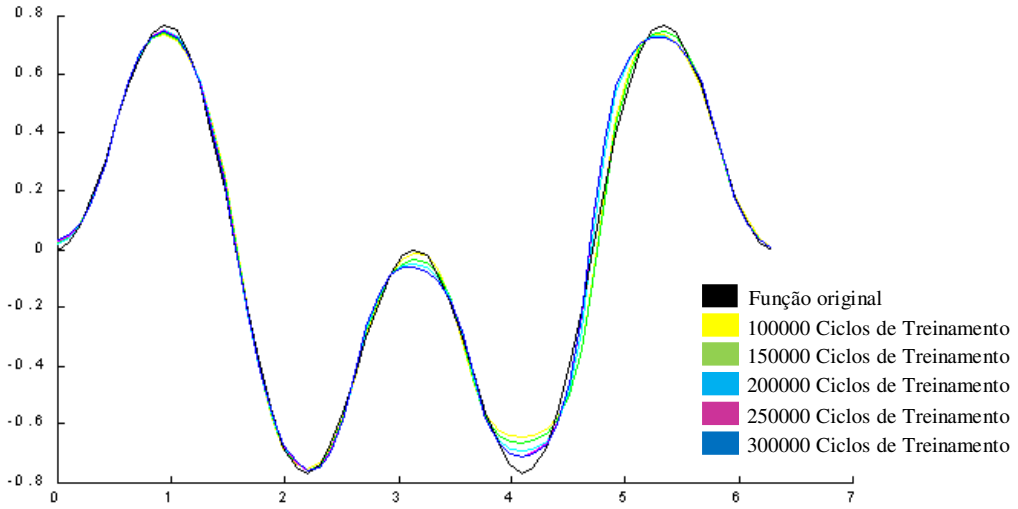

Figura 4 - Sobreposição das figuras geradas no treinamento de 100000 a 300000, com a taxa de aprendizado em 1.5.

Para verificar o poder de generalização da rede neural foi apresentado em sua entrada um conjunto de 241 valores, diferentes dos utilizados no treinamento. Comparando as respostas obtidas com as teóricas pode-se concluir que a rede tem uma satisfatória capacidade de generalização, apresentando sucesso em 97% dos casos.

CONCLUSÃO

Com o estudo realizado observa-se que com auxílio da rede MLP foi possível construir um sistema que substitui o modelamento de funções matemáticas, mostrando-se um aproximador com alta capacidade de generalização. Em vista destas propriedades apresentadas denota-se sua aplicabilidade em diferentes áreas como reconhecimento de padrões, processamento de sinais, controle entre outros.

Em grande parte dos sistemas de modelagem de um determinado fenômeno os sistemas clássicos de controle são utilizados para determinas as características do problema específico em questão. Existem muitos casos em que as informações disponíveis, proveniente do sistema, muitas vezes não apresentam um grau de confiabilidade ou características quantitativas suficientes para sua modelagem. O processo de modelagem se torna algo complexo, pois os sistemas clássicos necessitam de um conjunto de dados bem definidos para sua concretização, fazendo-se necessário o desenvolvimento de novas técnicas para os sistemas de controle.

Para estudos futuros recomenda-se uma evolução do caso apresentado na área de controle, através de dados experimentais a rede seria treinada e uma vez aprendendo a função não seria necessário à função de transferência para o modelamento.

REFERÊNCIAS

CAMPOS, M. M. de; SAITO, K. Sistemas inteligentes em controle e automação de processos – Rio de Janeiro: Editora Ciência Moderna Ltda., 2004.

FERNANDES, A. M. da R. Inteligência Artificial: noções gerais – Florianópolis: VisualBooks, 2005.

MACHADO, L. K. Rede Neural Artificial Embarcada em Robótica Móvel, 2012. Disponível: <http://www.scribd.com/doc/22506246/Rede-Neural-Artificial-Embarcada-em-Robotica-Movel> Acesso em: 16 nov. 2012

NASCIMENTO, J.; YONEYAMA, T. Inteligência artificial em controle e automação – São Paulo: Edgard Blücher: FAPESP, 2004.

SILVA, M. R. Compact yet efficient hardware architecture for multilayer-perceptron neural networks, 2011. Disponível: <http://www.scielo.br/scielo.php?pid=S0103-17592011000600010&script=sci_arttext> Acesso em: 14 out. 2012

TONSIG, L. S. Redes Neurais Artificiais Multicamadas e o Algoritmo Backpropagation, 2009. Disponível: <http://funk.on.br/esantos/doutorado/INTELIG%CANCIA%20ARTIFICIAL/T%C9CNICAS/REDES%20NEURAIS/Redes%20Neurais%20Artificiais%20Multicamadas/Backp.PDF> Acesso em: 15 out. 2012