

ESTUDO DE ALGORITMO MLP COMO APROXIMADOR DE FUNÇÃO

FRANCISCO ITALLO BRANDÃO RODRIGUES¹, RÔMULLO RANDELL MACEDO CARVALHO^{2*};
GUILHERME PEREIRA GONÇALVES³; KHALIL BELO PEREIRA BARBOSA⁴;

¹Graduando em Engenharia Elétrica, bolsista do PET, UFPI, Teresina-PI, itallobrandao7@hotmail.com

²Graduando em Engenharia Elétrica, bolsista do PET, UFPI, Teresina-PI, randellromullo@gmail.com

³Graduando em Engenharia Elétrica, UFPI, Teresina-PI, khalilbelo@gmail.com

⁴Graduando em Engenharia Elétrica, UFPI, Teresina-PI, brenohs77@hotmail.com

Apresentado no

Congresso Técnico Científico da Engenharia e da Agronomia – CONTECC'2016
29 de agosto a 1 de setembro de 2016 – Foz do Iguaçu, Brasil

RESUMO: Rede Neural Artificial, ou RNA, é uma categorização abrangente para algoritmos de programação que apresentam semelhança com o cérebro humano, com processamento de dados maciço paralelamente a partir de unidades simples, apresentando um crescente leque de aplicabilidade atualmente. Objetiva-se, neste estudo, analisar o desempenho de RNA supervisionada com aplicação direta em aproximação de funções. A RNA supervisionada utilizada foi o Perceptron Multicamada, ou MLP, treinada com algoritmo de retropropagação, pelo qual se pode realizar mapeamento não-linear de entrada-saída de natureza geral de forma prática. Para estudo, estipulou-se uma função real senoide somada a termo quadrático que, acrescida de um ruído branco gaussiano, representa os dados de saída da função para treino na MLP, caracterizando sinal ruidoso na entrada da rede, o que a torna mais robusta e prática. O resultado obtido a partir do processamento da rede é uma função próxima a função real estipulada não ruidosa, tão semelhante quanto melhor a configuração dos parâmetros atribuídos a rede, como a quantidade de épocas e de neurônios utilizados.

PALAVRAS-CHAVE: Rede neural artificial; Aproximação de funções; Perceptron multicamadas.

STUDY OF MLP ALGORITHM AS FUNCTION APPROXIMATOR

ABSTRACT: Artificial Neural network, or RNA, is a comprehensive categorization for programming algorithms that show similarity to the human brain, with massive parallel data processing from simple units, showing an increasing range of applicability. The goal is, in this study, analyze the performance supervised RNA with direct application in approximation of functions. The RNA supervised used was the Multilayer Perceptron, or MLP, trained with backpropagation algorithm, by which one can perform non-linear mapping of input-output of general nature in a practical way. To study, stipulated a real sine curve function plus the quadratic term that, plus a white Gaussian noise, represents the output data of the function for training at the MLP, featuring noisy signal at the entrance to the network, making it more robust and practical. The result obtained from the processing of the network is a close function the function stipulated noisy, not real so similar the better setting of the parameters assigned to the network, such as the amount of times and of neurons used.

KEYWORDS: Artificial neural network; Approximation of functions; Multilayer Perceptron.

INTRODUÇÃO

De modo bem semelhante ao que ocorre no cérebro humano, uma Rede Neural Artificial (RNA, denominada usualmente apenas como rede neural) apresenta um processamento de dados de acordo com determinados estímulos de entrada que culminam em uma ou mais saídas. A rede neural analisa essencialmente conhecimento experimental armazenado e o torna disponível para uso imediato ou posterior. De acordo com Haykin (2001), uma RNA pode ser caracterizada pelo processamento robusto e paralelamente distribuído em unidades de processamento simples denominadas neurônios.

Simplificadamente, redes neurais podem ser classificadas em supervisionadas e não supervisionadas; e para cada um dos dois tipos existem vários modelos disponíveis (BRAGA, 2007). A supervisão neste caso é atribuída ao fato da rede treinar segundo alguma disposição preliminar que a

supervisiona a um resultado; a não supervisão se dá quando os pesos se ajustam por conta própria, sem a necessidade dessa disposição.

Entre as redes supervisionadas, ressalta-se o *perceptron* multicamada (do inglês, *Multilayer Perceptron*, MLP), treinada com algoritmo de retropropagação de erro (*error back-propagation*), que se baseia em aprendizagem por correção de erro por propagação e retropropagação. Segundo Negnevitsky (2005), durante a aprendizagem, um conjunto de formação de padrões de entrada é apresentado a MLP, que calcula o padrão de saída; caso haja diferença entre a saída da rede e a desejada, os pesos são ajustados para reduzir este erro.

Em um *perceptron* de camada única, há apenas um peso para cada entrada e somente uma saída. No entanto, na MLP, existe uma camada de entrada, uma ou mais ocultas (ou intermediárias, ou escondidas) e uma de saída, com vários pesos em cada uma delas. A capacidade da MLP de aprender tarefas complexas se dá exatamente por conta dos neurônios da camada escondida, extraíndo progressivamente a cada época as características mais significativas dos padrões de entrada. Comumente, para aproximação de função, uma tarefa não tão complexa para uma ferramenta computacional poderosa como uma rede neural, utiliza-se apenas uma camada intermediária.

MATERIAIS E MÉTODOS

A implementação estrutural do algoritmo do *perceptron* multicamada neste artigo utiliza como ambiente de programação o *software* MATLAB, que apresenta uma interface amigável e funções predefinidas de cálculos operacionais não tão simples com desempenho considerável. Vale ressaltar ainda que ele permite manipulações de alto nível em programação para saída de gráficos e afins.

A função utilizada para aproximação neste artigo é apresentada em (1); esta apresenta três pontos locais máximos e dois pontos locais mínimos, o que torna a função caracteristicamente não tão simples de ser aproximada. Para treinamento da rede implementada, entretanto, acrescentou-se ainda na entrada do sinal o termo $e(x)$, como pode ser visto em (2); este termo representa um ruído branco gaussiano, que, em termos experimentais, é a modelagem de qualquer distúrbio adquirido durante uma aquisição em um sistema prático.

$$y = \sin(2x) + 0.25 \times (x - 3)^2 \quad (1)$$

$$y = \sin(2x) + 0.25 \times (x - 3)^2 + e(x) \quad (2)$$

O modelo de cada neurônio da rede inclui uma função de ativação não linear, com amplitude de -1 a 1 — essa função pode ter um padrão também de zero a um, mas se adota neste trabalho o primeiro intervalo por uma necessidade de se trabalhar com a região negativa da função aproximada. Para que a não linearidade da função de ativação seja essencialmente suave, adota-se neste estudo a função sigmoide, conforme (3).

$$f(x) = \frac{2}{1 + e^{-ax}} - 1 \quad (3)$$

Para confecção da rede utilizou-se por padrão uma camada de escondida, como usualmente ocorre para a aproximação de função; nesta camada, testou-se a rede para cinco, dez, vinte e cinquenta neurônios. Além disso, a rede foi comparada para 1000, 2500 e 6000 épocas; o objetivo é determinar para qual quantidade de neurônios escondidos e de épocas a rede melhor aproxima a função apresentada em (2).

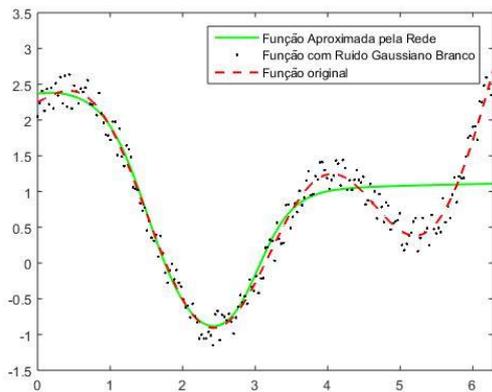
RESULTADOS E DISCUSSÃO

A Fig. 1 apresenta a saída e a entrada ruidosa da MLP, bem como a função real (não ruidosa) com 5 e 10 neurônios na camada oculta e com 1000, 2500 e 6000 épocas para a aproximação da função em (2); a Fig. 2 mostra os mesmos gráficos para 20 e 50 neurônios. Para 1000 épocas, independentemente da quantidade de neurônios, a saída da MLP não apresentou uma boa aproximação, sendo que, para 5, 10 ou 20 neurônios, isso se caracterizou nitidamente pela ausência de épocas suficientes, uma vez que o gráfico apresenta parcialmente um bom resultado, porém,

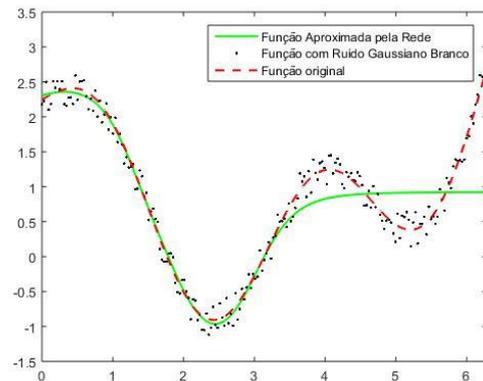
principalmente após o valor de π , o sinal de saída torna-se estável em um valor fixo (e não varia de acordo com a função real).

De forma geral, para 5, 10 ou 20 neurônios, o aumento de épocas melhorou a aproximação, sobretudo no exemplo de 6000 épocas. Entre esses três casos, o que melhor aproxima a função é com dez neurônios na cada oculta (Fig. 1a, 1b e 1c), que com 2500 épocas já apresenta boa aproximação. Com cinco neurônios (Fig. 1d, 1e e 1f), 2500 épocas ainda não é suficiente para aproximara adequadamente, mas apresenta uma melhora em relação a 1000; com 6000, tem-se já boa aproximação.

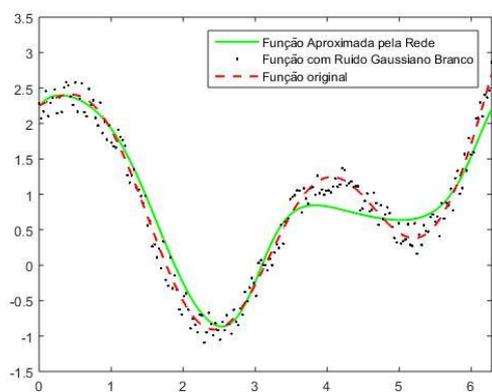
Figura 1 – Saída da MLP comparada à função real e à entrada ruidosa para (a) 5 neurônios e 1000 épocas, (b) 5 neurônios e 2500 épocas, (c) 5 neurônios e 6000 épocas, (d) 10 neurônios e 1000 épocas, (e) 10 neurônios e 2500 épocas e (f) 10 neurônios e 6000 épocas.



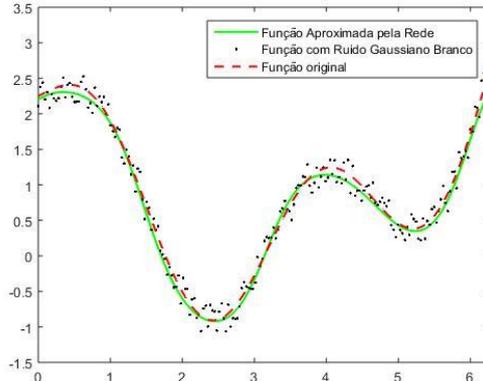
(a)



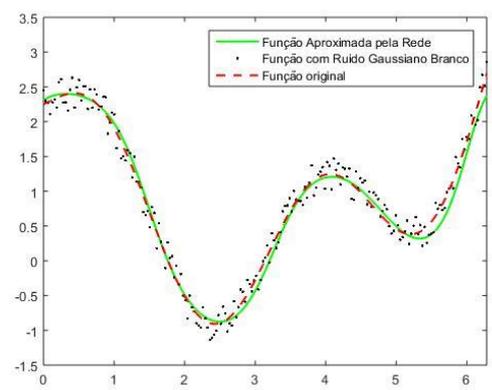
(d)



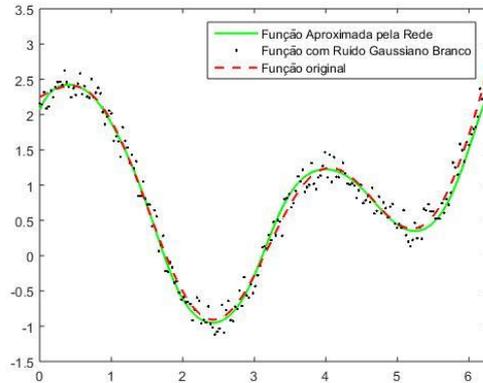
(b)



(e)



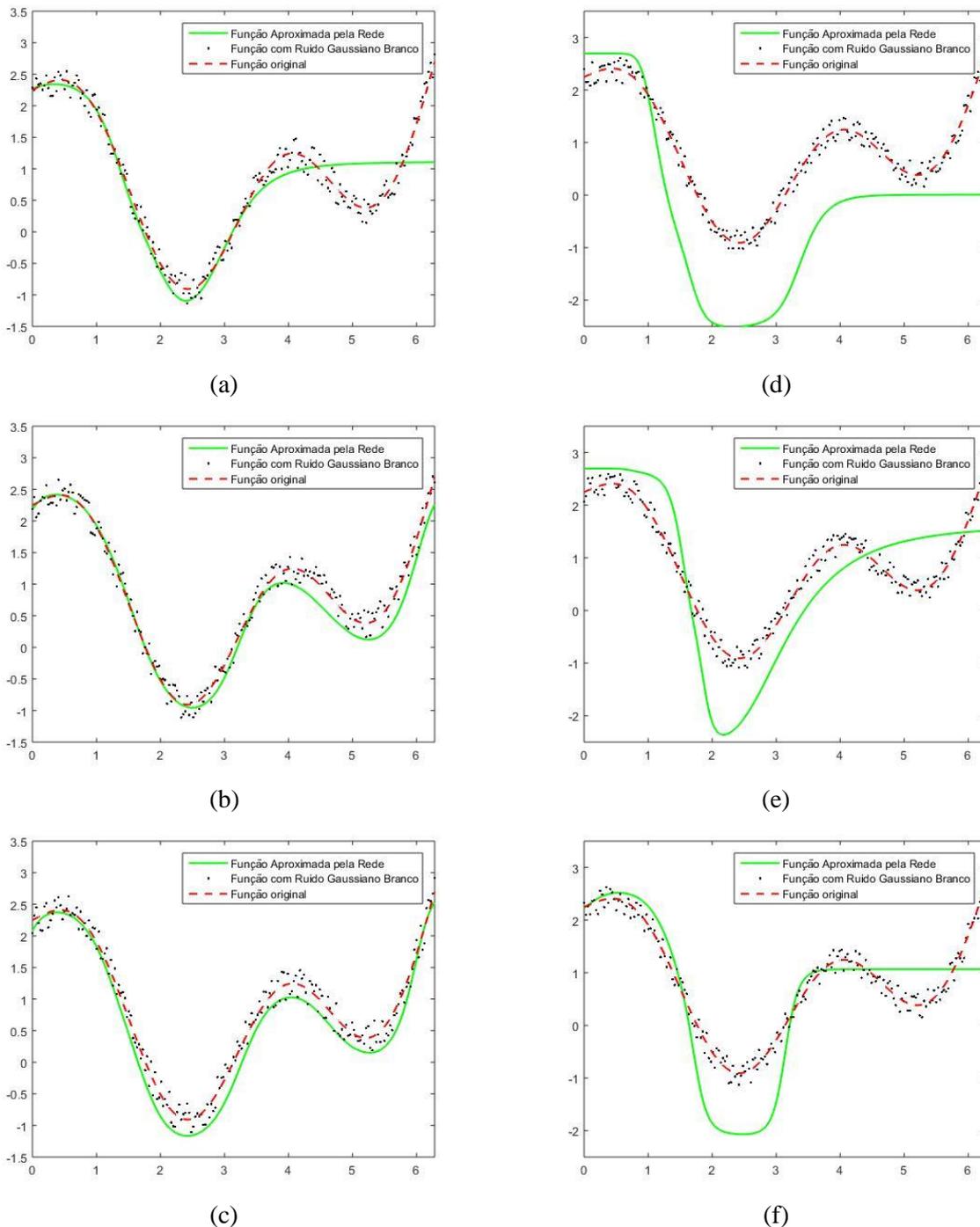
(c)



(f)

A rede MLP treinada com 20 neurônios na camada oculta (Figs. 2a, 2b e 2c) apresenta uma característica que já indicia o que ocorre para uma quantidade elevada de neurônios na camada escondida. O treinamento em poucas épocas (Fig. 2a) provoca o mesmo erro que para cinco ou dez neurônios; percebe-se que o aumento de épocas inicialmente melhora a aproximação (2500 épocas, Fig. 2b), porém, com novo aumento de épocas (6000, Fig. 2c), a aproximação é reduzida. Esse fenômeno, conhecido como *overfitting* é o treinamento excessivo da rede, ou seja, a quantidade de épocas é tão grande que a rede passa aprender o erro de retropropagação e não mais a função. Isso geralmente ocorre para um número bem maior de épocas, mas a quantidade demasiada de neurônios na camada oculta pode ter acelerado o processo.

Figura 2 – Saída da MLP comparada à função real e à entrada ruidosa para (a) 20 neurônios e 1000 épocas, (b) 20 neurônios e 2500 épocas, (c) 20 neurônios e 6000 épocas, (d) 50 neurônios e 1000 épocas, (e) 50 neurônios e 2500 épocas e (f) 50 neurônios e 6000 épocas.



Por fim, a rede MLP treinada com 50 neurônios na camada oculta (Figs. 2d, 2e e 2f) apresenta uma característica dissonante com todos os outros apresentados. Mesmo para 6000 épocas, a rede apresenta uma insuficiência de épocas para treino, aproximando a função, porém, com imprecisão grave. É provável que mesmo para valores superiores de épocas, essa quantidade de neurônios não apresente um comportamento ideal, uma vez que para 6000 épocas, a MLP com 20 neurônios já apresenta início de *overfitting*.

CONCLUSÕES

A rede supervisionada MLP implementada, quando parametrizada adequadamente, apresenta resposta satisfatória para a aproximação da função real apresentada em (1), mesmo com o sinal de entrada ruidoso, conforme (2). Esta função, no entanto, serviu apenas como exemplo assertivo da capacidade potencial da RNA em questão, uma vez que esta poderia aproximar uma gama de funções a disposição, desde que seus parâmetros de treinamento sejam devidamente ajustados.

Entre as quantidades de neurônios implementados, a MLP apresenta melhor precisão na saída com uma quantidade menor de neurônios na camada escondida (entre cinco e dez), necessitando de menos épocas para que o treinamento da rede culmine em uma resposta satisfatória quando comparada a função original. Por outro lado, o aumento de neurônios na camada escondida pode acelerar o processo de *overfitting* ou, se o aumento for ainda maior, resultar em uma aproximação com tamanha imprecisão que a função é descaracterizada.

AGRADECIMENTOS

Ao Programa de Educação Tutorial pela concessão de bolsa aos dois primeiros autores.

REFERÊNCIAS

- Braga, A. P.; Ferreira, A. C. P. L.; Ludermir, T. B. Redes neurais artificiais: teoria e aplicações. LTC Editora, 2007.
- Haykin, Simon. Redes neurais: princípios e prática. 2. ed. trad. Paulo Martins Engel. Porto Alegre: Bookman, 2001. pp. 27-282, 483-524.
- Negnevitsky, Michael. Artificial intelligence: a guide to intelligent systems. 2. Ed. King's Lynn: Pearson, 2007. pp. 175-185.
- Souza, L. G. M. Proposição e avaliação de algoritmos de filtragem adaptativa baseados na rede de kohonen. Fortaleza: UFCE, 2005.
- Souza, L. G. M. Modelos lineares locais para identificação de sistemas dinâmicos usando redes neurais competitivas. Fortaleza: UFCE, 2012.
- Von Zuben, F. J. Modelos paramétricos e não-paramétricos de redes neurais artificiais e aplicações. Campinas: FEE-UNICAMP, 1996.