

<https://eventos.utfpr.edu.br/sicite/sicite2019>

Redes neurais artificiais para classificação de dados

Artificial Neural Networks for data classification

RESUMO

Henrique Nazário Rocha
Nazario.utfpr@gmail.com
Universidade Tecnológica
Federal do Paraná, Ponta
Grossa, Paraná, Brasil

Hugo Valadares Siqueira
hugosiqueira@utfpr.edu.br
Universidade Tecnológica
Federal do Paraná, Ponta
Grossa, Paraná, Brasil

Este trabalho teve como objetivo aplicar Redes Neurais Artificiais para classificação de classes em dados reais. As arquiteturas utilizadas foram o Perceptron de Múltiplas Camadas (MLP), Redes Neurais com Funções de Base Radial (RBF), Máquina de Aprendizado Extremo (ELM) e a um modelo híbrido que envolve a RBF e a ELM. Foram usados três conjuntos de dados a fim de comparar o desempenho das mesmas. As amostras foram distribuídas em três conjuntos, treinamento, validação e testes, para que os resultados obtidos em teste após o treinamento da rede tenha valores consistentes, com as respectivas quantidades, 70%, 15% e 15% do valor total do conjunto de amostras. Os resultados mostram a viabilidade da proposta apresentada.

PALAVRAS-CHAVE: Redes Neurais, MLP, RBF. ELM.

ABSTRACT

This work had as objective to apply Artificial Neural Networks to real data classification. The architectures used were the Multilayer Perceptron (MLP), Radial Base Functions Networks (RBF), Extreme Learning Machines (ELM) and a hybrid model that involves the RBF and ELM. Three different databases were used to compare the performance of these networks. The samples were divided in three sets, training, validation and tests, with the respective quantities of 70%, 15%, and 15% of the total amount of the samples. The results showed the viability of the presented proposal.

KEYWORDS: Neural Networks. MLP. RBF. ELM.

Recebido: 19 ago. 2019.

Aprovado: 01 out. 2019.

Direito autorial: Este trabalho está licenciado sob os termos da Licença Creative Commons-Atribuição 4.0 Internacional.



INTRODUÇÃO

O cérebro humano é a máquina mais complexa que existe, tendo a capacidade de receber e processar estímulos em uma velocidade muito alta, usando processamento biológico de forma paralela. Isso só é possível graças aos neurônios, células que de forma organizada integram nosso sistema nervoso. Estas recebem, transmitem e processam estímulos em forma de pulsos nervosos, graças a um fenômeno eletroquímico que usam propriedades da própria membrana celular para produzir estímulos elétricos (excitatórios ou inibitórios) de forma coordenada. Existem algumas morfologias específicas neuronais, porém pode-se dizer que todo neurônio é composto basicamente por três partes principais: os dendritos, o corpo celular e o axônio (JUNQUEIRA, 2013).

Uma rede neural artificial (RNA) é um circuito, ou sistema, composto por uma grande quantidade de unidades simples inspiradas no sistema nervoso cerebral (NIGRIN, 1993). A representação do mesmo é dada por um modelo matemático simplificado do neurônio biológico, o qual possui pesos sinápticos ajustáveis com o intuito de emular o processo de aprendizagem. O neurônio artificial possui funções de receber, transmitir e processar sinais e, com a multiconexão de vários deles é possível a resolução de problemas pela troca de informações (HAYKIN, 1999). São capazes de reconhecer padrões de dados através do processo de treinamento e realizar generalizações baseada no aprendizado adquirido (DA SILVA, 2010).

REDES NEURAIS ARTIFICIAIS

O Perceptron de Múltiplas Camadas (MLP) é uma arquitetura de rede neural que contém no mínimo três camadas de neurônios: entrada, uma quantidade pré-determinada de ocultas ou intermediárias e uma de saída (HAYKIN, 1999). Por ter uma combinação de neurônios com partes não-lineares, as MLPs são redes também conhecidas por serem aplicáveis em diversos tipos de problemas, consideradas uma das arquiteturas mais versáteis. Algumas áreas são: aproximação universal, reconhecimento de padrões, identificação e controle de processos previsão de séries temporais e otimização de sistemas (DA SILVA, 2010).

De forma parecida com a MLP a arquitetura das Redes Neurais com Funções de Base Radial (RBF) possuem uma ampla variedade de aplicações (DA SILVA, 2010). A maneira como ela classifica ou realiza aproximações de padrões, porém, é distinta da anterior. Neste caso, normalmente consta apenas uma camada intermediária, na qual as funções de ativação são de alta dimensionalidade (HAYKIN, 1999).

A arquitetura de uma Máquina de Aprendizado Extremo (ELM) é bastante semelhante às MLP, mas com a diferença de que a camada intermediária não é treinada, ou seja, os pesos sinápticos são dados de forma aleatória e apenas os da camada de saída têm valores ajustados de forma determinística. Assim o esforço de processamento para o treinamento é bem menor (SIQUEIRA, 2012).

RESULTADOS E DISCUSSÃO

Para observar o comportamento das RNA em classificação foram utilizadas três bases de dados, sendo elas:

- Dados de telemarketing de um banco, os quais estão relacionados a campanhas de marketing ligados a uma instituição portuguesa (S. Moro, 2014);
- Base de dados Íris, que se refere a classificação de 3 espécies de flores, sendo que cada classe contém atributos das plantas como largura e comprimento da pétala e sépala (DUA, D., 2019);
- Base de dados de qualidade do vinho, que é classificada de acordo com variáveis físico-químicas (insumos) e sensoriais (a saída) (P. CORTEZ, 2009).

Os dados são divididos com a seguinte proporção: treinamento, com 70% da quantidade de amostras, validação com 15% e o conjunto teste os outros 15%. Para comparação do desempenho das redes apresentadas foram realizados testes com todas RNA citadas e com parâmetros diferentes como número de neurônios, quantidade de camadas e função de ativação. Neste último caso, utilizou-se as funções tangente hiperbólica Eq. (1) e logística Eq. (2):

$$\tanh = \frac{(e^t - e^{-t})}{(e^t + e^{-t})} \quad (1)$$

$$\log = \frac{1}{(1 + e^{-ta})} \quad (2)$$

As condições de parada para o processo de aprendizagem foram duas, erro quadrado médio entre a cada interação e quantidade de épocas em que o incremento da interação era insignificante para o resultado da rede. Para cada foram realizadas 10 rodadas, com uma precisão solicitada de 10^{-8} e uma variação de quantidade de neurônios semelhante.

Os resultados obtidos para cada base de dados são sumarizados nas Figuras 1, 2 e 3, nas quais Log define o uso da função logística, Tang da tangente hiperbólica, índice 1 para uma camada intermediária e 2 para duas. As Figuras apresentam o gráfico boxplot dos valores teste dos dados de acurácia de cada topologia de rede neural.

De acordo com os resultados podemos notar que algumas topologias de redes obtiveram uma acurácia próxima de 100% em todos os testes, como a maior parte das MLPs, RBFs e a proposta híbrida RBF+ELM. Nas bases de dados Iris e Vinho, é possível observar que a ELM com função logística obteve resultados inferiores. Na Figura 3 fica evidente que a topologia MLP teve uma taxa de acertos inferior as outras arquiteturas de rede, exceto aquela com tangente hiperbólica e 2 camadas intermediárias.

Figura 1 – Conjunto de dados Bank, com todas configurações de rede.

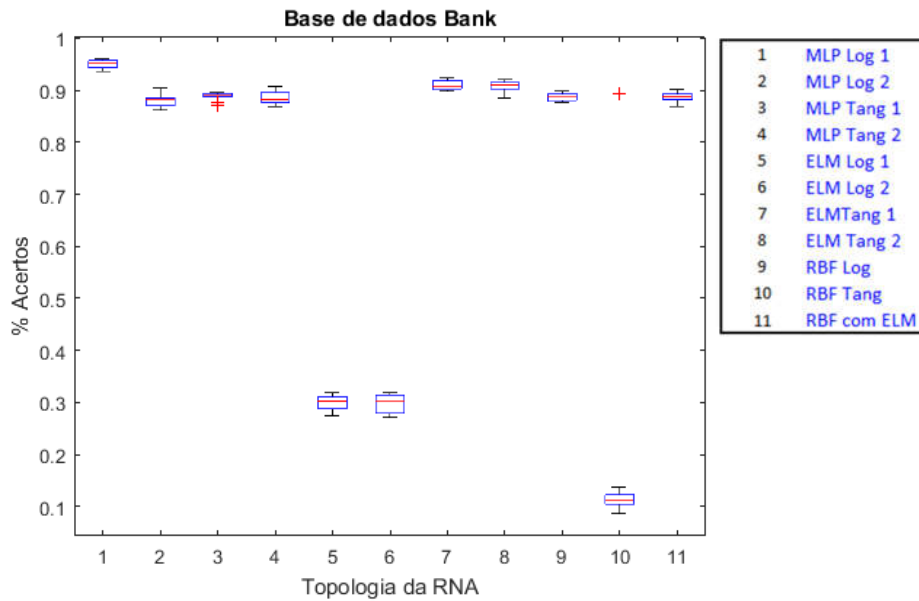


Figura 2 – Conjunto de dados Iris, com todas configurações de rede.

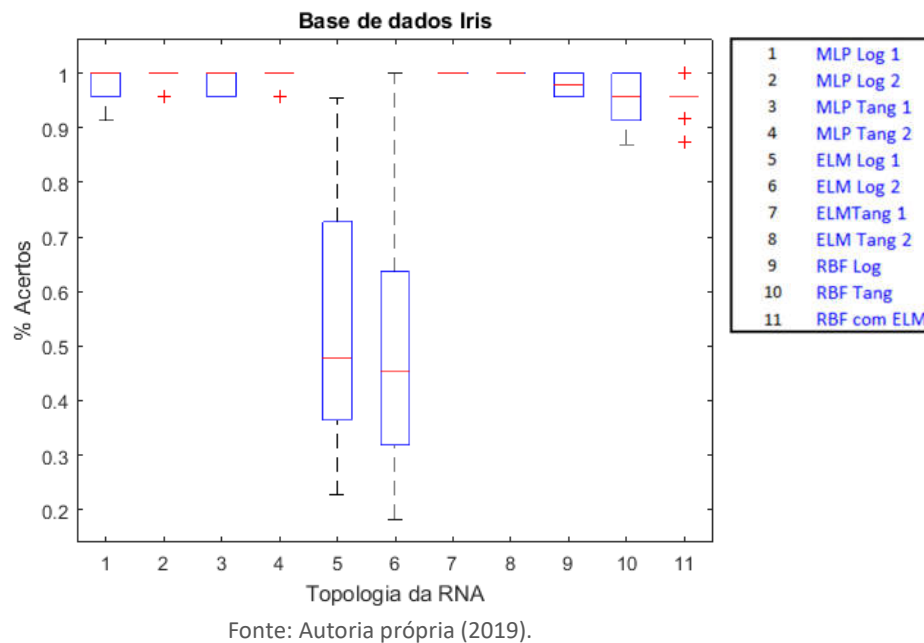
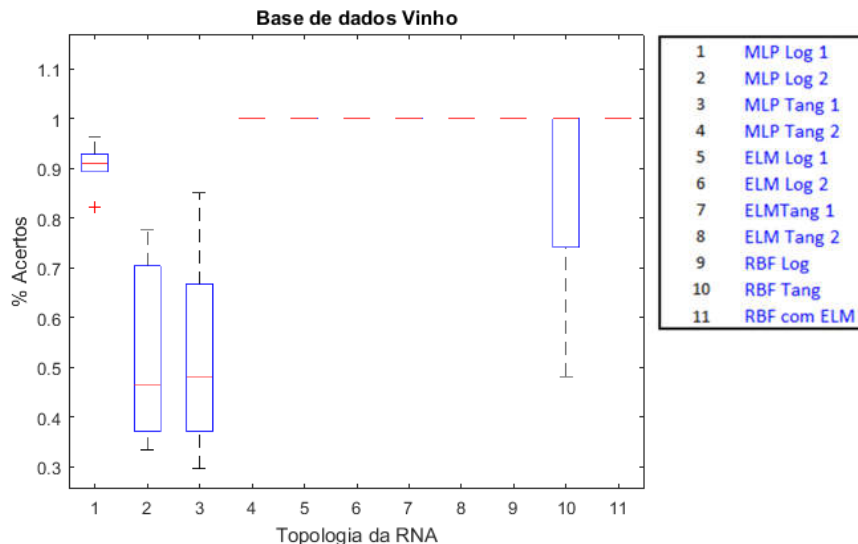


Figura 3 – Conjunto de dados Vinho, com todas configurações de rede.



Fonte: Autoria própria (2019).

CONCLUSÃO

O presente artigo apresentou a aplicação de Redes Neurais Artificiais (RNA) para classificação de padrões de três bancos de dados: Telemarketing de Banco, Iris e Qualidade de vinho. Foram implementadas e testadas o Perceptron de Múltiplas Camadas (MLP) com configurações de uma e duas camadas intermediárias, a Rede Neural com Função de Base Radial (RBF), Máquinas de Aprendizado Extremo (ELM) e a junção de duas redes, RBF com a ELM.

Os resultados mostram que a MLP, rede mais usada na literatura, teve um desempenho superior de classificação para as duas primeiras bases, mas foi relativamente inferior as demais na terceira. Neste caso, a ELM, RBF e a proposta híbrida chegaram a 100% de acerto, na maior parte dos casos. Ressalta-se que o custo computacional da ELM é muito inferior as demais.

Trabalhos futuros serão desenvolvidos no sentido de aplicar tais redes a outras bases de dados e a problemas de regressão e previsão.

AGRADECIMENTOS

Agradecemos ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo suporte financeiro, processo número 405580/2018-5.

REFERÊNCIAS

DA SILVA, I.N., SPATTI D., E FLAUZINO, R. (2010). **Redes neurais artificiais para engenharia e ciências aplicadas: curso prático**, Artliber Editora Ltda, São Paulo, SP, Brasil.

DUA, D. AND GRAFF, C. (2019). **UCI Machine Learning Repository** <http://archive.ics.uci.edu/ml>. Irvine, CA: University of California, School of Information and Computer Science.

HAYKIN, S. (1999) **NEURAL NETWORKS: A Comprehensive Foundation**. Prentice Hall, Upper Saddle River.

JUNQUEIRA, L. C.; CARNEIRO, J.; ABRAHAMSOHN, P. **Histologia básica: texto e atlas**. 12. ed. Rio de Janeiro: Guanabara Koogan, 2013.

NIGRIN, A. **Neural Networks for Pattern Recognition**, Cambridge, MA: The MIT Press, 1993.

P. CORTEZ, A. CERDEIRA, F. ALMEIDA, T. MATOS AND J. REIS. **Modeling wine preferences by data mining from physicochemical properties**. In Decision Support Systems, Elsevier.

SIQUEIRA, H. V., L. BOCCATO, R. ATTUX, AND C. LYRA FILHO. "Echo State Networks and Extreme Learning Machines: a Comparative Study on Seasonal Streamflow Series. Prediction." Lecture Notes in Computer Science, 2012a: 491-500.

S. Moro, P. Cortez and P. Rita. **A Data-Driven Approach to Predict the Success of Bank Telemarketing**. Decision Support Systems, Elsevier, 62:22-31, June 2014.