

O uso de algoritmos de otimização inspirados pela natureza na escolha de parâmetros de redes neurais artificiais

Igor Henrique Siriani
Fabricio Aparecido Breve
Unesp
Rio Claro, Brasil
igor.siriani@unesp.br

Resumo—Neste artigo, investigou-se o uso de algoritmos de otimização inspirados pela natureza para ajustar os parâmetros de redes neurais artificiais com o objetivo de melhorar seu desempenho. Para isso, comparamos o resultado obtido apenas pela rede neural com o resultado obtido pela rede parametrizada por um algoritmo Particle Swarm Optimization (PSO), por um algoritmo genético (GA) e por um algoritmo Imperialist Competitive Algorithm (ICA).

Área: “Inteligência Computacional”.

I. INTRODUÇÃO

As redes neurais artificiais (RNAs) podem ser interpretadas como um esquema de processamento capaz de armazenar conhecimento baseado em aprendizagem e disponibilizar este conhecimento para a aplicação em questão [1].

Um modelo básico de RNA possui diferentes componentes, como: Conjunto de sinapses, Integrador, Função de ativação e Bias [2]. Assim as RNAs são compostas por unidades de processamento chamadas neurônios, que se conectam entre si por meio de pesos sinápticos, além de outros componentes.

A estrutura e os parâmetros da rede neural influenciam diretamente no seu desempenho e na sua capacidade de generalização. Esses parâmetros podem ser determinados por meio de métodos heurísticos, experimentais ou analíticos e têm impacto significativo no desempenho das RNAs. Existem muitos estudos que apontam metodologias para determinação dos parâmetros [3], [4], [5], [6], [7]. No entanto, não existe um único método que escolha bons parâmetros para qualquer problema.

Uma alternativa para a escolha dos parâmetros da rede neural é o uso de algoritmos de otimização inspirados pela natureza. Esses algoritmos são técnicas meta-heurísticas que imitam processos naturais para encontrar soluções ótimas ou sub-ótimas para problemas difíceis ou complexos que exploram uma população de soluções candidatas e aplicam operadores estocásticos para gerar novas soluções, buscando melhorar a qualidade das soluções ao longo das iterações.

II. CONCEITOS E TÉCNICAS

O algoritmo de rede neural que usamos é uma rede neural artificial do tipo perceptron multicamadas (MLP), que é composta por uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Além disso usamos três algoritmos de otimização inspirados pela natureza para escolher o conjunto de pesos sinápticos da rede neural: Particle Swarm Optimization (PSO), Genetic Algorithm (GA) e Imperialist Competitive Algorithm (ICA). Esses algoritmos foram escolhidos por serem populares e eficientes em diversos problemas de otimização.

O problema que usamos para avaliar o desempenho dos algoritmos é a classificação das flores íris em três espécies: setosa, versicolor e virginica. Esse é um problema clássico e simples de aprendizado de máquina, que consiste em 150 amostras de flores, cada uma com quatro atributos: comprimento e largura da sépala e da pétala. O objetivo é construir um modelo capaz de prever a espécie de uma flor a partir dos seus atributos.

III. METODOLOGIA DE DESENVOLVIMENTO

Os valores de: número de camadas ocultas, o número de neurônios em cada camada oculta, a função de ativação dos neurônios e a taxa de aprendizagem foram mantidos fixos durante toda a comparação. Os valores utilizados nesses parâmetros são mostrados na “Tabela 1”.

TABELA 1 - VALORES DOS PARÂMETROS DA REDE NEURA.

Parâmetro	Valores
Número de camadas ocultas	1
Número de neurônios em cada camada oculta	4
Função de ativação dos neurônios	softmax
Número de neurônios na camada de entrada	4
Número de neurônios na camada de saída	3
Conjunto de treinamento	70%
Conjunto de validação	15%

Parâmetro	Valores
Conjunto de teste	15%
Épocas iniciais	25

IV. RESULTADOS OU RESULTADOS PRELIMINARES

Para avaliar os resultados dos algoritmos, usamos a métrica acurácia, que é a fração das previsões corretas da rede neural tendo sido medida antes da otimização, com as 25 iterações pré-estabelecidas e depois de terem sido aplicados os pesos sinápticos escolhidos pelos algoritmos de otimização inspirados pela natureza que estão em comparação nesse trabalho.

Para se buscar resultados mais precisos cada combinação de RNA e algoritmo de otimização foram executados cem vezes, das quais foram obtidas as acurácias mínimas, máximas, médias e o desvio padrão. Os resultados obtidos pelos algoritmos são mostrados na “Tabela 2”.

TABELA 2 - VALORES DOS PARÂMETROS DA REDE NEURAL.

Parâmetro	RNA	PSO	GA	ICA
Mínima	0,0870	0,4348	0,3913	0,3913
Máxima	1,0	1,0	1,0	1,0
Média	0,7890	0,8109	0,8509	0,8783
Desvio Padrão	0,1652	0,1556	0,1290	0,1319

O ICA ofereceu uma melhora de mais de 8,9 pontos percentuais em relação ao resultado da média do RNA. Além disso o PSO obteve mínimas melhores que o RNA base, representando um aumento de 500%, seguido do ICA e do GA que coincidentemente obtiveram os mesmos mínimos.

Isso indica que o ICA foi capaz de explorar melhor o espaço de busca do que os outros algoritmos e convergir para uma solução ótima ou sub-ótima. O GA e o PSO obtiveram resultados semelhantes, mas inferiores ao ICA, o que sugere que eles ficaram presos em ótimos locais ou não conseguiram explorar suficientemente o espaço de busca.

V. CONSIDERAÇÕES FINAIS

Os próximos passos para conclusão do projeto serão expandir as comparações para outros conjuntos de dados além do flores iris e expandir as comparações para outras arquiteturas de redes neurais além da MLP. Dessa forma, espera-se obter uma visão mais abrangente e diversificada sobre o desempenho dos algoritmos de aprendizado de

máquina para classificação de dados. Além disso, pretende-se realizar uma análise crítica dos resultados, identificando as vantagens e desvantagens de cada método, bem como as possíveis aplicações práticas e limitações.

REFERÊNCIAS

- [1] SPÖRL, Christiane; CASTRO, Emiliano; LUCHIARI, Aílton. Aplicação de redes neurais artificiais na construção de modelos de fragilidade ambiental. Revista Do Departamento De Geografia, v. 21, n. 1, p. 113-135, ago./2011. Disponível em: <https://doi.org/10.7154/RDG.2011.0021.0006>. Acesso em: 5 jun. 2023.
- [2] BRAGA, A. D. P.; CARVALHO, A. P. D. L. F. D.; LUDERMIR, Teresa Bernarda. Redes Neurais Artificiais: Teoria e Aplicações. 1. ed. Rio de Janeiro, RJ: Livros Técnicos e Científicos, 2000.
- [3] ISA, N. A. M.; MAMAT, W. M. F. W. Clustered-Hybrid Multilayer Perceptron network for pattern recognition application. Applied Soft Computing, Malaysia, v. 11, n. 1, p. 1457-1466, jan./2011. Disponível em: <https://doi.org/10.1016/j.asoc.2010.04.017>. Acesso em: 16 abr. 2023.
- [4] SIVARAM, G. S. V. S.; HERMANSKY, H. Sparse Multilayer Perceptron for Phoneme Recognition. IEEE, EUA, v. 20, n. 1, p. 23-29, mar./2011. Disponível em: <https://doi.org/10.1109/TASL.2011.2129510>. Acesso em: 17 abr. 2023
- [5] SAMAL, Aryapriyanka; PANDA, Jagyanseni; DAS, Niva. Performance Comparison of Single-Layer Perceptron and FLANN-Based Structure for Isolated Digit Recognition. Intelligent Computing, Communication and Devices, New Delhi, v. 308, n. 1, p. 237- 246, ago./2014. Disponível em: https://doi.org/10.1007/978-81-322-2012-1_25. Acesso em: 7 abr. 2023
- [6] KIM, Daehyon. Improving prediction performance of neural networks in pattern classification. International Journal of Computer Mathematics, v. 82, n. 4, p. 391- 399, mai./2005. Disponível em: <https://doi.org/10.1080/0020716042000301806>. Acesso em: 13 abr. 2023.
- [7] WANG, Xiaotong; CHANG, Chih-chen; DU, Fang. Achieving a More Robust Neural Network Model for Control of a MR Damper by Signal Sensitivity Analysis. Neural Computing & Applications, v. 10, n. 1, p. 330-338, abr./2002. Disponível em: <https://doi.org/10.1007/s005210200005>. Acesso em: 28 abr. 2023.