

Otimização do pré-processamento de Redes Neurais através de Nuvem de Partículas e Algoritmo Genético

João M. Fernandes¹, Paulo H. Siqueira²
PPGMNE/UFPR, Curitiba, PR

Uma das primeiras etapas para a execução de uma rede neural é o pré-processamento dos dados de entrada e de saída. Assim, [1] comenta sobre a importância dessa etapa, destacando que existem diversas possibilidades de se realizar esses processos tornando possível melhorar consideravelmente a taxa de acerto da rede neural final. [2] também indicou a necessidade de uma análise mais profunda desse tema, explicando que a adoção de técnicas simples pode prejudicar significativamente a performance da rede neural. Portanto, o objetivo do trabalho é compreender a variação nos resultados devido ao pré-processamento e desenvolver uma metodologia para otimizar o tratamento de dados.

A primeira etapa foi limitar quais técnicas de pré-processamento seriam estudadas. Como consequência, são consideradas 3 etapas principais para o tratamento dos dados: tratamento dos valores faltantes; normalização dos dados e redução da dimensão das variáveis de entrada. Para os valores faltantes, foi definida a possibilidade de remover as amostras, substituir os valores pela média ou mediana, média de grupos de amostras baseada em outras variáveis ou utilizar o KNNinpute [3]. Já para a normalização dos dados, optou-se por adicionar as opções de normalização mínima-máxima, normalização Z-score ou normalização do máximo absoluto. Por fim, foi colocada a possibilidade de se realizar ou não uma análise de componentes principais (ACP).

Para prosseguir com o estudo, foi selecionada uma base de dados, com o objetivo de se prever a quantidade de anéis de abalones por meio de suas propriedades físicas (medidas e pesos). A base de dados contém 8 variáveis de entradas, das quais 5 contêm dados faltantes, entre 10 e 26% do total de amostras. Dessa forma, optou-se pelos esquemas de pré-processamento apresentados na Tabela 1.

Tabela 1: Possibilidades de tratamentos de dados.

Variável ou Técnica	Modificações	Possibilidades
Length		5
Diameter	Média, Mediana, Média baseada na Idade,	5
Height	Média baseada no Sexo ou KNNinpute	5
Whole Weight	Remover, Média, Mediana, Média baseada no	5
Viscera Weight	Sexo ou KNNinpute	5
ACP	Não realizar ou 2,3 componentes	5
Normalização	Média, Mediana, Média baseada na Idade,	5
	Média baseada no Sexo ou KNNinpute	
Total		28125

Com as possibilidades levantadas, foram executadas as 28125 possibilidades de rede neural e

¹fernandesjoaomauricio@gmail.com

²paulohs@ufpr.br

foi construído um histograma com base na Raiz do Erro Quadrático Médio (REQM) encontrado, apresentado na Figura 1 a).

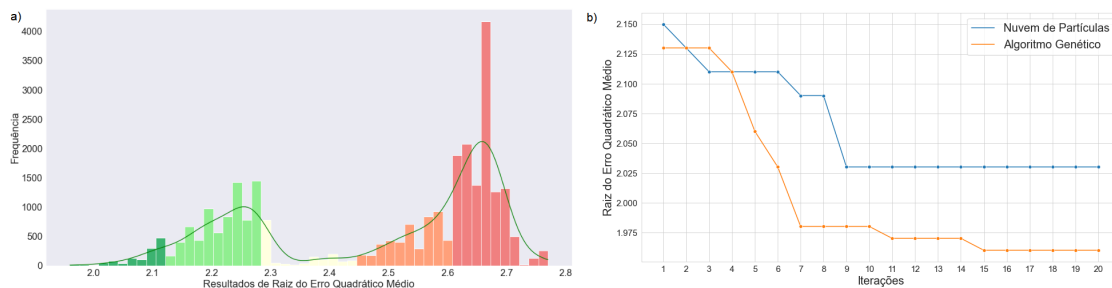


Figura 1: a) Histograma dos Resultados. b) Convergência dos algoritmos de otimização. Fonte: Os autores.

A Figura 1 a) torna evidente a necessidade de estudos acerca de técnicas para seleção do pré-processamento. Portanto, foram testadas duas técnicas para essa seleção. Inicialmente foi utilizada uma otimização por Nuvem de Partículas, adaptada para considerar as probabilidades de realizar cada pré-processamento [4], permitindo a sua utilização em variáveis discretas. Além disso, foi realizada uma otimização com Algoritmo Genético por ranqueamento. Ambas as técnicas são comparadas na Figura 1 b).

Nos resultados fica evidente que ambos os algoritmos funcionaram bem e conseguiram convergir para resultados interessantes de REQM. Entretanto, é importante destacar que o Algoritmo Genético conseguiu evoluir mais rapidamente, atingindo a melhor possibilidade de pré-processamento disponível.

Dessa forma, pode-se concluir que ambos os algoritmos são capazes de encontrar soluções interessantes para o tratamento de dados, reduzindo a necessidade de testes e melhorando a performance da rede neural. Além disso, para ambas as metodologias, foi necessária acrescentar um fator de aleatoriedade significativo. Para a Nuvem de Partículas, a seleção aleatória entre as probabilidades e para o Algoritmo Genético uma mutação em quase 50% dos indivíduos de cada geração. Portanto, ambos os algoritmos podem ser considerados técnicas viáveis para a seleção adequada dos melhores pré-processamentos disponíveis.

Referências

- [1] S. Asadi, E. Hadavandi, F. Mehmanpazir e M. M. Nakhostin. “Hybridization of evolutionary Levenberg-Marquardt neural networks and data pre-processing for stock market prediction”. Em: **Knowledge-Based Systems** 35 (2012). DOI: 10.1016/j.knsys.2012.05.003.
- [2] J. Han e S. Kang. “Dynamic imputation for improved training of neural network with missing values”. Em: **Expert Systems with Applications** 194 (2022). DOI: 10.1016/j.eswa.2022.116508.
- [3] O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, T. Hastie, Tibshirani R., D. Botstein e R. B. Altman. “Missing value estimation methods for DNA microarrays”. Em: **Bioinformatics** 17 (2001), pp. 520–525.
- [4] S. Strasser, R. Goodman, J. Sheppard e S. Butcher. “A new discrete Particle Swarm Optimization algorithm”. Em: **GECCO 2016 - Proceedings of the 2016 Genetic and Evolutionary Computation Conference**. 2016. DOI: 10.1145/2908812.2908935.