Neural Networks vs. Regression: A Comparative Analysis in Medical Data Processing

Minodora ANDOR[a] and Gheorghe Ioan MIHALAŞ[b,*]

[a] Department of Internal Medicine 1, Victor Babeş University of Medicine and Pharmacy, Eftimie Murgu Sq., no. 2, 300041 Timişoara, Romania.
[b] Center for Biological Systems Modeling and Data Analysis, Victor Babeş University of Medicine and Pharmacy, Eftimie Murgu Sq., no. 2, 300041, Timişoara, Romania.
E-mails: andorminodora@gmail.com, mihalas@gmail.com

* Author to whom correspondence should be addressed;

Abstract
Background and Aim: The increasing adoption of artificial intelligence (AI) in medical research offered alternative methods for medical data processing. This study evaluated comparatively the predictive performance of feedforward neural networks (FFNN) regression versus classical statistical regression analysis in estimating the risk of post-COVID-19 type 2 diabetes based on metabolic factors. The primary objective was to assess the applicability, advantages, and limitations of these approaches when applied to relatively small medical datasets. *Materials and Methods:* We started with the analysis of a small dataset - 130 patient records with metabolic parameters [1]. The risk of post-COVID-19 type 2 diabetes (glycaemia at 4 and at 12 months post-COVID as function of metabolic parameters) was predicted using both linear regression and FFNN. The regression model followed standard statistical guidelines, while the FFNN required optimization of hyperparameters, including the number of layers, activation functions, learning rate, and optimization algorithms. We extended the study using simulated data to further compare logistic regression (a data set of 300 patients) with neural networks. *Results:* The classical regression models demonstrated stable performance with clear interpretability, offering well-defined coefficients and statistical significance measures. However, FFNN did not yield superior predictive accuracy, and its performance varied significantly depending on the choice of hyperparameters. The optimization process for NN required extensive trial and error, as no universal guidelines for parameter selection were applicable in this context. *Discussion.* Our findings highlight a real challenge in medical AI applications for data processing: when dealing with small datasets, neural networks do not necessarily outperform classical methods. Regression provided robust results with minimal computational effort, while FFNN required complex tuning without a clear performance advantage. The use of simulated data revealed that NN might be more effective in larger datasets with potential non-linear patterns, but limited interpretability. *Conclusion:* AI-based models are, indeed, recommended for data processing of large and/or unstructured complex medical data sets. However, as a conclusion of this study, regression models proved to be a more practical and reliable choice for small-scale medical predictions. Future work should explore hybrid models that combine interpretability with non-linear modeling capacity to optimize predictive accuracy in clinical settings.

Keywords: Neural Networks; Regression Analysis; Medical Data Processing; Predictive Modeling.

Reference
1. Andor M, Man DE, Nistor DC, Buda V, Dragan S. The Influence of COVID-19 in Glycemic Control: Predictive Value of Inflammation and Metabolic Parameters. Biomedicines 2024;12:2642. https://doi.org/10.3390/biomedicines12112642.