REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

PREDIÇÃO DA ATIVIDADE ANTIOXIDANTE DE FRUTAS ATRAVÉS DO USO DE REDE NEURAL ARTIFICIAL

PREDICTION OF ANTIOXIDANT ACTIVITY OF FRUITS THROUGH THE USE OF ARTIFICIAL NEURAL NETWORK

Hélio Júnior Alvarenga Godinho
Engenheiro de Alimentos, Universidade Federal de Lavras, Brasil
E-mail: hjagodinho@hotmail.com

Danilo José Machado de Abreu
Doutor em Microbiologia Agrícola, Instituto de Ciências Naturais (ICN), Universidade Federal de Lavras, Brasil
E-mail: danilo.mabreu@gmail.com

Diogo Nunes Carvalho
Engenheiro de Controle e Automação, Departamento de Computação (DC), Universidade Federal de São Carlos, Brasil
E-mail: diogonunescarv@gmail.com

Estela Corrêa De Azevedo
Mestre em Ciência dos Alimentos, Escola de Ciências Agrárias de Lavras (ESAL), Universidade Federal de Lavras, Brasil
E-mail: estela.azevedo1@estudante.ufla.br

Adeilson Carvalho
Mestre em Engenharia Agrícola, Escola de Ciências Agrárias de Lavras (ESAL), Universidade Federal de Lavras, Brasil
E-mail: adeilson@ufla.br

Elisângela Elena Nunes Carvalho
Doutora em Ciência dos Alimentos, Escola de Ciências Agrárias de Lavras (ESAL), Universidade Federal de Lavras, Brasil
E-mail: elisangelacarvalho@ufla.br

Mário Sérgio Lorenço
Doutor em Engenharia de Biomateriais, Instituto de Ciência, Tecnologia e Inovação (ICTIN), Universidade Federal de Lavras, Brasil
E-mail: mariolorenco@ufla.br

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

Resumo

Os compostos fenólicos são reconhecidos por sua expressiva bioatividade, atuando na proteção vegetal contra estresses abióticos e proporcionando benefícios terapêuticos à saúde humana. Embora o interesse por alimentos funcionais tenha impulsionado a demanda por análises de qualidade, os métodos analíticos convencionais para determinar a capacidade antioxidante são onerosos e demandam elevado tempo operacional. Nesse contexto, a aplicação de ferramentas de Inteligência Artificial surge como uma alternativa estratégica para a otimização de processos. O presente estudo objetivou predizer a atividade antioxidante de diferentes matrizes frutíferas utilizando Redes Neurais Artificiais (RNA). A base de dados foi estruturada a partir de dez artigos científicos, dos quais foram extraídos perfis fenólicos e valores de atividade antioxidante (DPPH). Para o desenvolvimento do modelo, utilizou-se a arquitetura *Multilayer Perceptron* (MLP) com 13 variáveis de entrada e três camadas ocultas. Os dados foram normalizados pelo método Min-Max. O banco de dados completo (N = 22, incluindo juçara) foi avaliado por validação cruzada Leave-One-Out (LOO-CV), com comparação de desempenho frente à Ridge Regression e Regressão PLS. Diferentes funções de ativação e algoritmos de otimização foram testados para avaliar a acurácia preditiva. Os resultados demonstraram que a combinação da função de ativação tangente hiperbólica (*tanh*) com o otimizador gradiente descendente estocástico (SGD) apresentou o melhor desempenho, atingindo um Erro Percentual Absoluto Médio (MAPE) de $3{,}93\times10^{-6}$ e Erro Quadrático Médio (MSE) de $2{,}28\times10^{-10}$ no ajuste ao conjunto de treinamento. A validação cruzada Leave-One-Out (LOO-CV, N = 22) revelou que a Ridge Regression apresentou melhor capacidade de generalização (R^2 = 0,763; RMSE = 784,6) do que a MLP (R^2 = −0,119), indicando que a complexidade da topologia (20-15-10) não é justificada pelo tamanho amostral atual. O estudo posiciona-se como investigação exploratória de prova de conceito, demonstrando a viabilidade da predição de atividade antioxidante a partir de perfis fenólicos por aprendizado de máquina, com perspectiva de validação prospectiva em conjuntos de dados mais amplos.

Palavras-chave: compostos fenólicos; alimentos funcionais; aprendizado de máquina.

Abstract

Phenolic compounds are recognized for their significant bioactivity, playing a role in plant protection against abiotic stresses and providing therapeutic benefits to human health. Although interest in functional foods has driven the demand for quality analysis, conventional analytical methods for determining antioxidant capacity are costly and require high operational time. In this context, the application of Artificial Intelligence tools emerges as a strategic alternative for process optimization. The present study aimed to predict the antioxidant activity of different fruit matrices using Artificial Neural Networks (ANNs). The database was structured from ten scientific articles, from which phenolic profiles and antioxidant activity values (DPPH) were extracted, totaling 22 samples across eight fruit species. For model development, a Multilayer Perceptron (MLP) architecture was used with

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

13 input variables and three hidden layers. Data were normalized using the Min-Max method and evaluated via Leave-One-Out Cross-Validation (LOO-CV, N = 22), with benchmark comparison against Ridge Regression and PLS. Different activation functions and optimization algorithms were tested to evaluate predictive accuracy. The results demonstrated that the combination of the hyperbolic tangent activation function (*tanh*) with the Stochastic Gradient Descent (SGD) optimizer yielded the best performance, achieving a Mean Absolute Percentage Error (MAPE) of 3.93×10^{-6} and a Mean Squared Error (MSE) of 2.28×10^{-10} on the training set. Leave-One-Out Cross-Validation (LOO-CV, N = 22) revealed that Ridge Regression achieved better generalization ($R^2 = 0.763$; RMSE = 784.6) than the MLP ($R^2 = -0.119$), indicating that the (20-15-10) topology complexity is not justified by the current sample size. This study is positioned as an exploratory proof-of-concept investigation demonstrating the feasibility of antioxidant activity prediction from phenolic profiles using machine learning, with prospective external validation as the recommended next step.

Keywords: phenolic compounds; functional foods; machine learning.

1. Introduction

Brazil ranks among the world's leading fruit producers, driven by an exceptional biodiversity that yields matrices rich in bioactive compounds (FARIAS; SANCHES; PETRUS, 2023). Interest in these foods has grown exponentially, grounded in the ability of certain molecules to mitigate oxidative stress in the human body. Free radicals, natural byproducts of cellular metabolism or external sources such as pollution and radiation, are highly reactive species capable of causing irreversible damage to proteins, lipids, and DNA, and have been associated with the development of degenerative diseases and premature aging (CHANDIMALI *et al.*, 2025).

In this context, phenolic compounds emerge as the primary antioxidant defense agents. These substances, which include flavonoids, anthocyanins, and phenolic acids, are essential secondary metabolites for plant physiology, providing protection against ultraviolet radiation and pathogen attack. In the human diet, their efficacy stems from a chemical structure rich in hydroxyl groups, which enable free radical scavenging through the donation of electrons or hydrogen atoms, thereby interrupting harmful oxidation chain reactions (BARROS DORES *et al.*, 2026; LV *et al.*, 2021; ZHANG *et al.*, 2022).

Despite the recognized importance of these compounds, their quantification

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

and the determination of total antioxidant capacity, most commonly performed via the DPPH (2,2-diphenyl-1-picrylhydrazyl) assay, present significant operational bottlenecks. Conventional methods are inherently dependent on high-cost reagents, require time-consuming extraction protocols, and generate chemical waste, all of which constrain the agility of industrial quality control and the feasibility of large-scale screening (SADOWSKA-BARTOSZ; BARTOSZ, 2022).

As an alternative to these limitations, the integration of Industry 4.0 technologies, specifically Artificial Neural Networks (ANNs), represents a paradigm shift in food science. ANNs are computational mathematical models inspired by human neurological architecture, capable of processing large volumes of data and identifying complex nonlinear patterns. Unlike simple linear statistical models, multilayer networks (Multilayer Perceptron) can learn the relationship between the chemical profile of a fruit (input variables) and its corresponding antioxidant activity (output variable), enabling fast and accurate predictions without the need for exhaustive new laboratory analyses (GORBACHEV *et al.*, 2022).

In recent literature, machine learning models have been successfully applied to the prediction of physicochemical and bioactive properties of food matrices. Studies have demonstrated that MLP architectures are capable of modeling nonlinear relationships between chromatographic profiles and biological responses with high accuracy, outperforming multiple linear regression models in chemically complex systems (JUNG *et al.*, 2024; LIAKOS *et al.*, 2025). Nevertheless, the application of these tools to the specific prediction of antioxidant capacity in Brazilian tropical fruits, matrices characterized by wide phenolic variability, remains largely unexplored, which provides the rationale for the present investigation.

Given the complexity of Brazilian fruit matrices and the need for accessible predictive tools, this study aims to develop and validate an ANN model capable of estimating antioxidant activity from the phenolic profile of different fruit species. Through this approach, we seek to provide a low-cost, technologically efficient method for the functional characterization of food products.

2. Literature Review

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

2.1 Fruit production and consumption

World fruit production has shown continuous growth over the past decades. According to FAO/FAOSTAT data, global fruit production reached 951.51 million tons in 2023, harvested across approximately 68.10 million hectares, representing a consistent linear growth trend over the previous ten (PEREIRA GONZATTO; SCHERER SANTOS, 2025).

On the world stage, Brazil has consolidated itself as the third largest fruit producer in the world, with an annual production of 43 million tons and a Gross Production Value of R$ 75 billion in 2023, trailing only China and India (PETRY *et al.*, 2025). Most of this production is aimed at the domestic consumer market, with exports concentrated in varieties such as mango, melon, papaya, grapes, lemons and limes, reaching demanding markets in Europe, North America and Asia. The Brazilian fruit sector is highly diversified, with oranges leading national production, followed by bananas, açaí, grapes and pineapples, and with family farming playing a central role in the supply of the domestic market (PETRY *et al.*, 2025).

Despite Brazil's prominent position as a global fruit producer, adequate fruit and vegetable consumption among the adult population remains below recommended levels. Data from the Surveillance of Risk and Protective Factors for Chronic Diseases by Telephone Survey (VIGITEL), covering the period from 2008 to 2023, indicate that recommended fruit and vegetable consumption has shown a declining trend since 2015, with a reduction of approximately 0.52 percentage points per year in the proportion of adults meeting the World Health Organization (WHO) recommendation of at least 400 grams per day (VEIGA *et al.,* 2025). This gap between production capacity and actual consumption patterns underscores the importance of research initiatives aimed at enhancing the functional characterization and quality assessment of fruit matrices, providing scientific support for the promotion of these foods as sources of bioactive compounds beneficial to human health.

2.2 Functional foods

Since the 1960s, a nutritional transition has been underway in Brazil, characterized by a shift in the population's dietary habits through the reduction in

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

consumption of fresh and minimally processed foods and the increase in ultra-processed products with high concentrations of sodium, sugars, and fats (LOUZADA *et al.*, 2023). Data from the Brazilian Household Budget Survey indicate that the contribution of ultra-processed foods to total dietary energy intake increased by 44% between 2003 and 2018, representing a public health concern with significant implications for the epidemiological and nutritional profile of the Brazilian population (LOUZADA *et al.*, 2023).

In this context, the search for functional foods has intensified. Functional foods can be defined as foods or ingredients that offer health benefits beyond their basic nutritional requirements, not being intended to cure diseases, but contributing to the promotion and maintenance of health through the action of bioactive compounds on physiological processes (DEBRI *et al.*, 2026; DIXIT *et al.*, 2023).

Fruits, recognized sources of vitamins, minerals, dietary fiber, and phytochemicals, occupy a prominent position in this category. The protective effect attributed to these foods has been largely associated with the presence of polyphenols, particularly flavonoids, phenolic acids, and anthocyanins, which possess potent antioxidant action (RUDRAPAL *et al.*, 2022). Epidemiological evidence strongly suggests an inverse association between diets rich in fruits and vegetables and the incidence of cardiovascular diseases, certain cancers, metabolic syndrome, and neurodegenerative disorders (RUDRAPAL *et al.*, 2022).

The antioxidant capacity of polyphenols is primarily attributed to their reducing properties, which depend fundamentally on the number and position of hydroxyl groups present in the molecule. These groups enable the donation of electrons or hydrogen atoms to neutralize reactive oxygen species, interrupting harmful oxidation chain reactions. The intensity of this antioxidant action varies significantly among different classes of polyphenols as a function of their structural characteristics (LV *et al.*, 2021; RUDRAPAL *et al.*, 2022).

2.3 Phenolic compounds

Phenolic compounds are secondary metabolites widely distributed in nature, with over 8,000 phenolic molecules already identified in plants. This large and

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

structurally diverse group is present in a variety of fruits, vegetables, cereals, and processed products, where they function as pigments responsible for coloration, UV radiation protectors, and defense agents against pathogens and environmental stressors. From a nutritional standpoint, these compounds act as antioxidants not only through their capacity to donate hydrogen atoms or electrons to neutralize reactive oxygen species, but also through the stability of the intermediate radicals formed, which inhibit the propagation of oxidative chain reactions and protect food lipids from peroxidation (LV *et al.*, 2021; ZHANG *et al.*, 2022).

Phenolic compounds share a common structural feature, namely an aromatic ring bearing one or more hydroxyl substituents. Based on their carbon skeleton and structural complexity, this class is broadly divided into two major groups: flavonoids and non-flavonoids (SHAHRAJABIAN; SUN, 2023). Flavonoids represent the largest subgroup, comprising compounds with a C6–C3–C6 diphenylpropane skeleton, and include several subclasses widely distributed in fruits and vegetables, such as flavonols, flavones, flavanones, flavan-3-ols (catechins), anthocyanins, isoflavones, and chalcones (SUN; SHAHRAJABIAN, 2023; ZHANG *et al.*, 2022).

In the non-flavonoid class, phenolic acids are the most representative compounds, subdivided into derivatives of hydroxybenzoic acids (C6–C1 skeleton) and hydroxycinnamic acids (C6–C3 skeleton). Common hydroxybenzoic acids include gallic, protocatechuic, vanillic, and syringic acids, while the main hydroxycinnamic acids are caffeic, p-coumaric, ferulic, and sinapic acids (ROCCHETTI *et al.*, 2022; SUN; SHAHRAJABIAN, 2023). The antioxidant activity of phenolic acids is closely related to the number and position of hydroxyl groups on the aromatic ring, as well as to the structural relationship between the carboxyl group and the phenyl ring, which determines the electron delocalization capacity and, consequently, the radical scavenging efficiency of these molecules (PLATZER *et al.*, 2022).

2.4 Artificial neural networks

Artificial Neural Networks (ANNs) are computational models inspired by the biological structure and functioning of the human brain, consisting of interconnected

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

processing units called artificial neurons that perform mathematical operations. These models are designed to simulate cognitive capabilities such as learning, generalization, pattern recognition, and abstraction when subjected to training with representative data (MADHIARASAN; LOUZAZNI, 2022).

ANNs are particularly suited for input-output mapping of nonlinear systems and for performing parallel processing tasks, making them especially valuable for simulating complex systems and generating coherent responses for data patterns not previously encountered during training. Unlike conventional computer programs based on explicit instruction sets, ANNs acquire knowledge through examples: the learning algorithm adjusts the network's adaptive parameters, known as synaptic weights, to generalize relationships present in the training data (ALZUBAIDI *et al.*, 2021; MADHIARASAN; LOUZAZNI, 2022).

An ANN is fundamentally characterized by two elemental components: its architecture and its learning algorithm. The architecture defines the structural organization of neurons and their connections, including the number of layers and the connectivity pattern between them. The learning algorithm governs how the synaptic weights are updated during training to minimize prediction error. At the individual neuron level, processing typically consists of a weighted linear combination of inputs followed by the application of a nonlinear activation function, which introduces the representational capacity necessary to model complex, real-world relationships (ALZUBAIDI *et al.*, 2021; GOODFELLOW; BENGIO; COURVILLE, 2017).

2.5 Prediction neural networks

The applicability of artificial neural networks has expanded significantly in recent decades, driven by the growing availability of computational tools and the increasing demand for methods capable of modeling complex nonlinear relationships. ANNs have been successfully applied across a wide range of domains, including regression, classification, function approximation, pattern recognition, optimization, and multivariate data analysis, demonstrating their versatility in both scientific and industrial contexts (ALZUBAIDI *et al.*, 2021; MADHIARASAN; LOUZAZNI, 2022).

In the field of food science and engineering, ANNs have emerged as

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

particularly powerful tools for modeling nonlinear processes and predicting quality-related responses. Applications include the modeling of extraction, drying, fermentation, thermal processing, and quality evaluation operations, in all of which the networks demonstrated high predictive accuracy while requiring no prior mechanistic knowledge of the underlying system (BHAGYA RAJ; DASH, 2022). Neural networks have also proven suitable for real-time applications, given their capacity to generate fast and accurate responses once trained, making them attractive for integration into automated quality control systems in the food industry (BHAGYA RAJ; DASH, 2022; LIAKOS *et al.*, 2025).

The main advantages of ANNs for predictive applications include fault tolerance, self-adaptation capacity, and the ability to resolve complex problems without requiring the explicit definition of rules or mathematical models. These characteristics make them particularly valuable in situations involving classification, pattern association, function approximation, and decision-making in dynamic environments where it would be difficult or impractical to formulate precise analytical models (ALZUBAIDI *et al.*, 2021; GOODFELLOW; BENGIO; COURVILLE, 2017).

2.6 Multilayer Perceptron (MLP) neural networks

A Multilayer Perceptron (MLP) is a class of feedforward artificial neural network composed of at least three layers of interconnected nodes: an input layer, one or more hidden layers, and an output layer. In this architecture, information propagates unidirectionally from the input layer, through the hidden layers, to the output layer, where the final prediction or classification is obtained (ALZUBAIDI *et al.*, 2021; GOODFELLOW; BENGIO; COURVILLE, 2017).

MLPs are characterized by full connectivity between adjacent layers: each neuron in a given layer receives input from all neurons in the preceding layer and transmits its output to all neurons in the subsequent layer. This fully connected structure distinguishes the standard MLP from partially connected architectures, in which some inter-layer connections are absent. The hidden layers constitute the computational core of the network, enabling it to extract and refine increasingly abstract representations of the input data. The depth of the network, defined by the

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

number of hidden layers, determines its capacity to model intricate nonlinear relationships that would be inaccessible to linear models (ALZUBAIDI *et al.*, 2021).

One of the fundamental properties of the MLP is its ability to solve nonlinearly separable problems, which is only possible when the neurons in the hidden layers employ nonlinear activation functions. Without such nonlinearities, successive linear transformations would collapse into a single equivalent linear mapping, regardless of network depth, severely limiting its representational power. This principle is formally supported by the universal approximation theorem, which establishes that an MLP with a single hidden layer and a nonlinear activation function is capable of approximating any continuous function on a compact subset of $\mathbb{R}^n$ to arbitrary accuracy, given a sufficient number of neurons (GOODFELLOW; BENGIO; COURVILLE, 2017; MADHIARASAN; LOUZAZNI, 2022).

2.7 Activation functions

The activation function is the nonlinear transformation applied to the output of each neuron in a neural network. It determines whether a neuron should be activated based on the weighted sum of its inputs, introducing the nonlinearity that allows the network to model complex, non-linearly separable relationships in data. Without activation functions, successive layers of a neural network would collapse into a single linear transformation regardless of depth, severely limiting its representational capacity (DUBEY; SINGH; CHAUDHURI, 2022). The most commonly used activation functions in ANNs are described below.

Logistic (Sigmoid): The logistic sigmoid function was for many years one of the primary activation functions used in neural networks. It maps any real-valued input to the interval (0, 1), making it useful for probability-based outputs and binary classification tasks. Compared to the original binary threshold function (Heaviside step function), the sigmoid presents the advantage of being continuously differentiable throughout its domain, which enables parameter updating through the backpropagation algorithm. However, a known limitation of the sigmoid is the vanishing gradient problem: for very large or very small input values, the function saturates and its gradient approaches zero, slowing or halting learning in deep

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

networks (DUBEY; SINGH; CHAUDHURI, 2022). Its equation is:

$$\sigma(X) = \frac{1}{1 + e^{-x}} \quad (1)$$

Hyperbolic tangent (*tanh*): A recognized disadvantage of the sigmoid function is that it is not zero-centered, producing outputs exclusively in the positive range (0, 1). This characteristic can introduce unfavorable dynamics during gradient-based optimization by causing consistently unidirectional weight updates. The hyperbolic tangent function addresses this limitation by mapping inputs to the range (−1, 1), centering the output around zero and generally enabling faster convergence compared to the sigmoid, while maintaining analogous differentiability properties (DUBEY; SINGH; CHAUDHURI, 2022). Its equation is:

$$\sigma(x) = \tanh x = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \quad (2)$$

Rectified Linear Unit (*ReLU*): The ReLU function became widely adopted in modern deep learning architectures due to its simplicity and computational efficiency. It returns zero for negative inputs and the identity value for positive inputs, effectively introducing sparsity in network activations. Key advantages of ReLU include its non-saturating behavior for positive inputs, which mitigates the vanishing gradient problem, and its compatibility with the stochastic gradient descent optimization method, enabling faster training convergence compared to sigmoid-based functions (DUBEY; SINGH; CHAUDHURI, 2022; GOODFELLOW; BENGIO; COURVILLE, 2017). Its equation is:

$$\sigma\,(x) = max\ (0, x)\ \ (3)$$

3. Methodology

The methodology followed a structured workflow comprising four main stages: database compilation, data preprocessing, neural network architecture design, and statistical validation.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

3.1 Database Compilation

Data were compiled from previous studies that employed High-Performance Liquid Chromatography (HPLC) for phenolic profile characterization, combined with the DPPH assay for antioxidant activity determination. The dataset consisted of ten scientific articles covering jabuticaba, guabiju, jambolão, yellow guava, acerola, and blackberry matrices at different ripening stages (AIDI WANNES; MARZOUK, 2013; BAI; ZHANG; REN, 2013; BETTA *et al.*, 2018; FENG *et al.*, 2017; NASCIMENTO *et al.*, 2018; PARIKH; PATEL, 2017; SCHULZ *et al.*, 2015, 2020; SERAGLIO *et al.*, 2018; WANG *et al.*, 2018). Prior to model development, the antioxidant activity response variable (DPPH) was harmonized across all source studies to a common unit of expression: milligrams of ascorbic acid equivalents per 100 g of dry matter (mg AAE 100 g^{-1} DM). Source studies reporting DPPH in alternative units were converted using the respective reported equivalence factors before inclusion in the dataset. Phenolic compound concentrations were similarly verified for unit consistency (mg g^{-1} fresh matter) across all matrices. Initially, 38 compounds were identified and subjected to screening based on their frequency of occurrence across the matrices and their recognized antioxidant potential in the literature, yielding 21 candidate variables (RUBERTO; BARATTA, 2000). Additionally, samples from juçara fruit (*Euterpe edulis* Martius) present a biochemically distinct profile dominated by anthocyanins and other compounds not quantified by the 13 target markers, which systematically underrepresents their antioxidant-relevant composition. These samples were retained in the full dataset (N = 22) to enable a comprehensive assessment of model performance, including across chemically heterogeneous matrices. The dataset was then refined to 13 input variables, selected by frequency of occurrence across all 22 matrices (≥ 22.7%), ensuring sufficient representation for model training across the LOO-CV iterations. The complete variable selection workflow is summarized in Figure 1.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

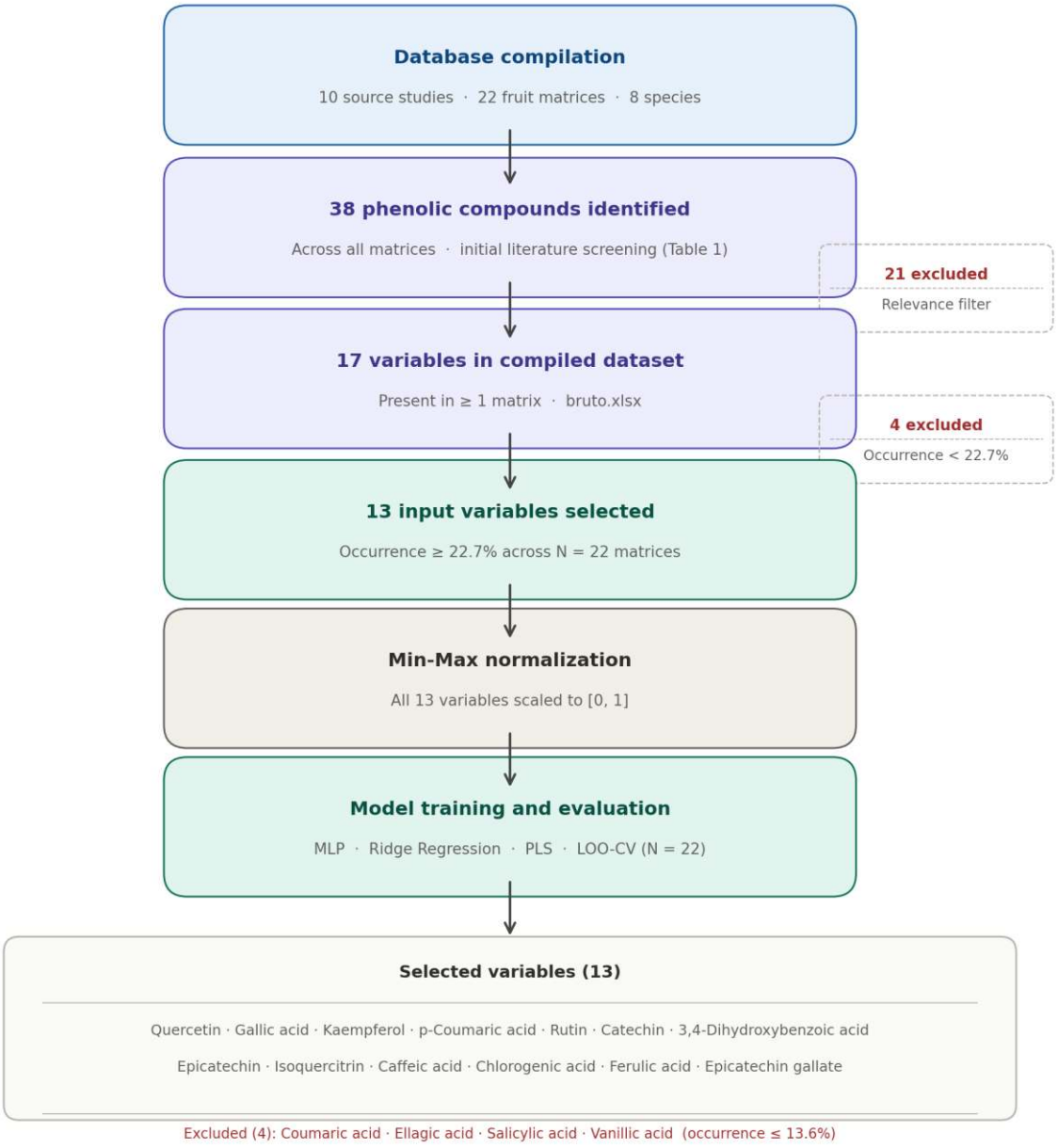

Figure 1. Variable selection workflow: from 38 phenolic compounds identified across 10 source studies to the 13 input variables retained for model training (occurrence ≥ 22.7% across N = 22 matrices).

3.2 Preprocessing and Normalization

To ensure comparability across matrices, all concentrations were standardized to mg/g of fresh matter. Data were normalized using the Min-Max method, scaling each variable to the range [0, 1] in order to preserve the original distribution without

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

distortion from outlying values. Given the small sample size (N = 22), Leave-One-Out Cross-Validation (LOO-CV) was adopted as the primary validation strategy, as recommended for datasets with fewer than 30 observations (GOODFELLOW; BENGIO; COURVILLE, 2017). In each LOO iteration, one sample is held out as the test set while the remaining N−1 samples are used for training, yielding N independent predictions without data leakage. Benchmark comparison was additionally performed against Ridge Regression and Partial Least Squares Regression (PLS, 2 and 3 components) using the same LOO-CV framework.

3.3 Artificial Neural Network Architecture

Implementation was carried out in Python (v3.6.9) on the Google Colab platform, using the scikit-learn library. The adopted architecture was the Multilayer Perceptron (MLP) in its MLPRegressor configuration. The network design (Figure 2) consisted of:

- Input Layer: 13 neurons corresponding to the selected phenolic compounds.
- Hidden Layers: Three intermediate layers with a topology of (20, 15, 10) neurons, defined empirically through systematic testing of alternative configurations, totaling 45 processing units.
- Output Layer: 1 neuron dedicated to predicting the antioxidant activity value. The maximum number of epochs was set to 5,000 (max_iter=5000; n_iter_no_change=5000), a value adopted to ensure convergence without imposing premature constraints on learning. The complete set of MLPRegressor parameters was: hidden_layer_sizes = (20, 15, 10); activation = 'tanh' (best-performing configuration); solver = 'sgd'; max_iter = 5,000; n_iter_no_change = 5,000; learning_rate_init = 0.001; tol = 1×10^{-4} (scikit-learn default); random_state = 42; alpha = 0.0001 (L2 penalty, default). Input features were scaled using MinMaxScaler (range [0, 1]); the DPPH response variable was used in its original measurement scale. The following activation functions were tested in combination with the Adam and Stochastic Gradient Descent (SGD) optimizers, totaling six evaluated configurations:
    - Logistic (Sigmoid): Historically prominent in ANN models, this function is

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

governed by Equation (1). Its main advantage lies in a non-zero derivative across its entire domain, which facilitates learning through the backpropagation algorithm (DUBEY; SINGH; CHAUDHURI, 2022).

- Hyperbolic Tangent (*tanh*): Unlike the sigmoid, the *tanh* function is zero-centered, as represented by Equation (2). This property is desirable to avoid unfavorable dynamics during gradient minimization, while retaining the differentiability properties of the logistic function (DUBEY; SINGH; CHAUDHURI, 2022).
- Rectified Linear Unit (*ReLU*): Defined by Equation (3), *ReLU* stands out for its low computational cost and its ability to accelerate training via the stochastic gradient descent method (DUBEY; SINGH; CHAUDHURI, 2022; GOODFELLOW; BENGIO; COURVILLE, 2017).

The optimal configuration was identified based on LOO-CV performance metrics (R^2, RMSE, MAE), evaluated on hold-out samples not used during model training.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

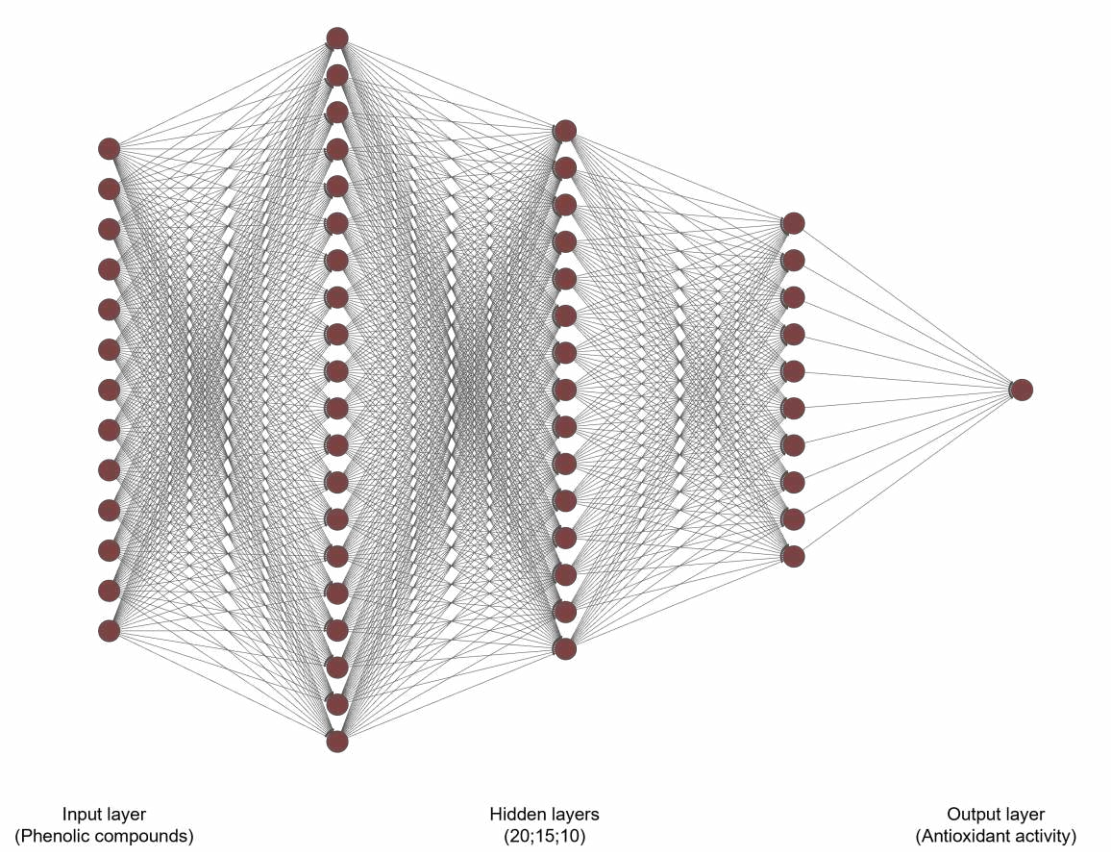

Figure 2. Architecture of the Multilayer Perceptron (MLP) neural network used for antioxidant activity prediction, comprising an input layer with 13 neurons (phenolic compounds), three hidden layers with a (20-15-10) neuron topology, and an output layer with 1 neuron (antioxidant activity determined by the DPPH assay).

Model performance under LOO-CV was assessed using R^2, RMSE, and MAE. The MAPE and MSE, as defined by Equations (4) and (5), were additionally calculated for descriptive comparison of in-sample fit across activation function configurations and are reported in Table 4 for reference only.

$$MAPE = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \quad (4)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (5)$$

Where A_t and y_i represent the observed real values, while F_t and $\hat{y}_i$ represent the values predicted by the neural network.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

4. Results and Discussion

The thirty-eight compounds found in the studies (AIDI WANNES; MARZOUK, 2013; BAI; ZHANG; REN, 2013; BETTA *et al.*, 2018; FENG *et al.*, 2017; NASCIMENTO *et al.*, 2018; PARIKH; PATEL, 2017; SCHULZ *et al.*, 2015, 2020; SERAGLIO *et al.*, 2018; WANG *et al.*, 2018) can be observed in Table 1 below.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

Table 1. Phenolic compounds identified across the fruit matrices analyzed and subjected to screening for selection of the predictive model input variables.

Phenolic Compound	Phenolic Compound
3,4-Dihydroxybenzoic acid	Epicatechin
Benzoic acid	Phlorizin
Caffeic acid	Epicatechin gallate
Chlorogenic acid	Hispidulin
Coumaric acid	Isoquercitrin
Ellagic acid	Isorhamnetin
Ferulic acid	Kaempferol
Gallic acid	L-Epicatechin
p-Coumaric acid	Luteolin
Protocatechuic acid	Myricetin
Salicylic acid	Naringenin
Sinapic acid	Pinobanksin
Syringic acid	Quercetin
Vanillic acid	Quercitrin
Coniferyl aldehyde	Rutin
Aromadendrin	Sinapaldehyde
Catechol	Syringaldehyde
Catechin	Taxifolin
Chrysin	Vanillin

The initial data screening revealed that the inclusion of samples from juçara fruit (*Euterpe edulis* Martius) present a phenolic profile dominated by anthocyanins and other compounds not captured by the 13 target markers used in this study, resulting in a mean total quantified phenolic content of 1.93 mg/g FM — compared to 9,631.15 mg/g FM for the remaining matrices (Mann-Whitney U = 0; $p < 0.0001$) — and a mean DPPH activity of 10.17 vs. 1,569.92 mg AAE 100 g^{-1} (Mann-Whitney U =

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

3; p = 0.0001). Despite this biochemical divergence, juçara samples were retained in the full dataset (N = 22) to provide a complete assessment of model performance across heterogeneous matrices. The LOO-CV framework treats juçara samples as independent test observations in each iteration, allowing their prediction error to be quantified transparently rather than removed. A sensitivity analysis comparing LOO-CV performance with (N = 22) and without juçara (N = 15) is presented in the Limitations section. The dataset was refined to 13 input variables, selected by frequency of occurrence across the 22 fruit matrices (≥ 22.7%). These variables comprise compounds of recognized antioxidant relevance, such as Gallic Acid, Quercetin, and Rutin, whose chemical structures are illustrated in Figure 3.

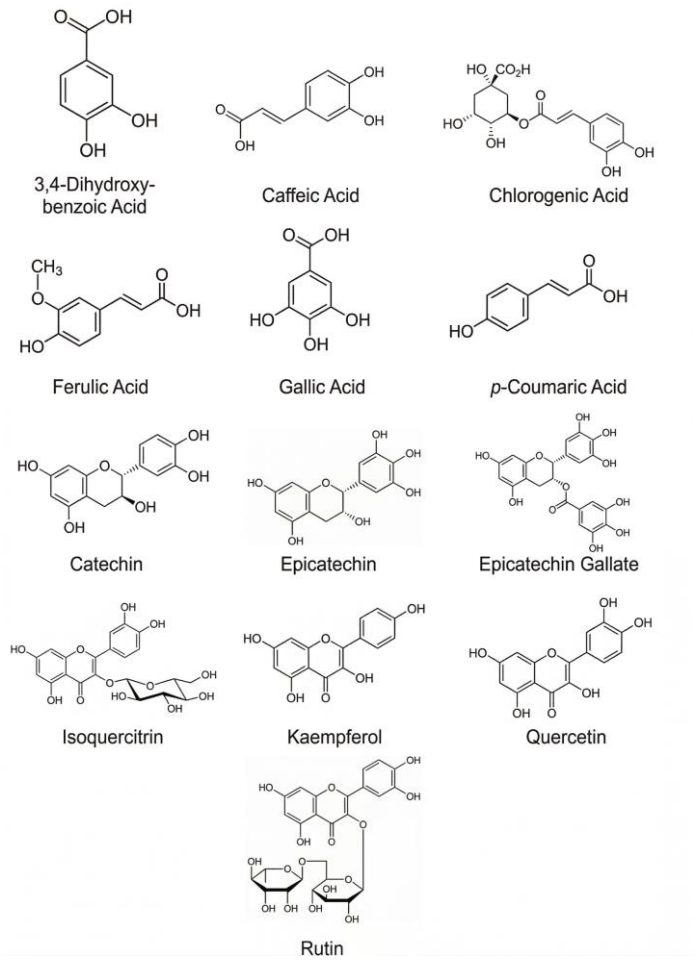

Figure 3. Chemical structures of the 13 phenolic compounds selected as input variables for the antioxidant activity predictive model.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

The dataset containing the phenolic compound concentrations can be found in Table 2, along with the antioxidant activity values used as output data in the neural network. Additionally, the normalized dataset used as network input is presented in Table 3.

Table 2. Phenolic compound concentrations (mg/g fresh matter) in the analyzed fruit matrices and corresponding antioxidant activity values (DPPH) used as the predictive model output variable.

Fruit	ADB[1]	CA[2]	CHA[3]	FA[4]	GA[5]	pCA[6]	CAT[7]	EC[8]	ECG[9]	IQ[10]	KF[11]	QU[12]	RU[13]	DPPH[14]
Jabuticaba (unripe)	1709.15	70.4	348.71	502.26	7680.83	531.3	12.84	0	0	677.46	78.56	6795.33	0	5066.35
Jabuticaba (ripe)	1299.74	25.41	193.27	199.74	4164.51	358.47	0	0	0	0	33.48	5211.41	0	4662.09
Guabiju (unripe)	127.13	41.97	11.08	128.87	7780.34	170.15	194.81	36.22	1577.83	1417.87	1079.66	78	84.81	3105.99
Guabiju (ripe)	82.88	25.13	5.86	0	4379.98	111.51	48.37	11.22	971.72	2160.35	901.11	8240.85	0	2569.28
Jambolão (unripe)	30.64	0	0	0	21494.1	70.36	0	0	25.75	0	0	98.45	0	2761.32
Jambolão (ripe)	36.03	0	0	0	11203.1	50.07	0	0	0	13.16	0	44.72	0	2760.5
Yellow guava (unripe)	64.06	0	18.13	0	532.61	53.69	11424.2	0	32.99	3158.31	32.85	1833.67	0	19.69
Yellow guava (mid-ripe)	88.81	0	11.87	124.56	781.93	90.31	10880.74	0	0	2622.13	228.97	1118.72	0	20.01
Yellow guava (ripe)	75.6	54.28	8.74	79.28	1542.01	63	12795.5	22.34	0	1194.6	32.7	1038.77	0	18.46
Blackberry (stage I)	4.3	0	0	0	6.8	0	101	381.4	0	189.1	0.9	53.5	128.1	248.8
Blackberry (stage II)	12	0	0	0	6.8	0	3.9	8.6	0	191.8	1.8	10.2	96.3	331.1
Acerola (stage I)	3.1	9.5	0	12.7	3.8	9.8	8	7	0	431.9	16.1	120.8	0	112.4
Acerola (stage II)	3.9	7.3	0	6.2	4	7.9	6.4	6.3	0	331.6	23.8	119.8	0	96.33

Footnotes: [1] 3,4-Dihydroxybenzoic acid | [2] Caffeic acid | [3] Chlorogenic acid | [4] Ferulic acid | [5] Gallic acid | [6] p-Coumaric acid | [7] Catechin | [8] Epicatechin | [9] Epicatechin gallate | [10] Isoquercitrin | [11] Kaempferol | [12] Quercetin | [13] Rutin | [14] Antioxidant activity. Table shown for the 13 matrices used in the original model development. The full dataset (N = 22) used for LOO-CV evaluation is available from the corresponding author upon request

Table 3. Min-Max normalized dataset used as input for the neural network antioxidant activity prediction model.

ADB[1]	CA[2]	CHA[3]	FA[4]	GA[5]	pCA[6]	CAT[7]	EC[8]	ECG[9]	IQ[10]	KF[11]	QU[12]	RU[13]	DPPH[14]
0.0171	0.0007	0.0035	0.0050	0.0768	0.0053	0.0001	0.0000	0.0000	0.0068	0.0008	0.0680	0.0000	50.6635
0.0130	0.0003	0.0019	0.0020	0.0416	0.0036	0.0000	0.0000	0.0000	0.0000	0.0003	0.0521	0.0000	46.6209
0.0013	0.0004	0.0001	0.0013	0.0778	0.0017	0.0019	0.0004	0.0158	0.0142	0.0108	0.0008	0.0008	31.0599
0.0008	0.0003	0.0001	0.0000	0.0438	0.0011	0.0005	0.0001	0.0097	0.0216	0.0090	0.0824	0.0000	25.6928
0.0003	0.0000	0.0000	0.0000	0.2149	0.0007	0.0000	0.0000	0.0003	0.0000	0.0000	0.0010	0.0000	27.6132
0.0004	0.0000	0.0000	0.0000	0.1120	0.0005	0.0000	0.0000	0.0000	0.0001	0.0000	0.0004	0.0000	27.6050
0.0006	0.0000	0.0002	0.0000	0.0053	0.0005	0.1142	0.0000	0.0003	0.0316	0.0003	0.0183	0.0000	19.6900
0.0009	0.0000	0.0001	0.0012	0.0078	0.0009	0.1088	0.0000	0.0000	0.0262	0.0023	0.0112	0.0000	20.0100
0.0008	0.0005	0.0001	0.0008	0.0154	0.0006	0.1280	0.0002	0.0000	0.0119	0.0003	0.0104	0.0000	18.4600
0.0043	0.0000	0.0000	0.0000	0.0068	0.0000	0.1010	0.3814	0.0000	0.1891	0.0009	0.0535	0.1281	24.8000
0.0120	0.0000	0.0000	0.0000	0.0068	0.0000	0.0039	0.0086	0.0000	0.1918	0.0018	0.0102	0.0963	33.1000
0.0031	0.0095	0.0000	0.0127	0.0038	0.0098	0.0000	0.0070	0.0000	0.4319	0.0161	0.1208	0.0000	112.4000
0.0039	0.0073	0.0000	0.0062	0.0040	0.0079	0.0064	0.0063	0.0000	0.3316	0.0238	0.1198	0.0000	96.3300

Footnotes: [1] 3,4-Dihydroxybenzoic acid | [2] Caffeic acid | [3] Chlorogenic acid | [4] Ferulic acid | [5] Gallic acid | [6] p-Coumaric acid | [7] Catechin | [8] Epicatechin | [9] Epicatechin gallate | [10] Isoquercitrin | [11] Kaempferol | [12] Quercetin | [13] Rutin | [14] Antioxidant activity. Table shown for the 13 matrices used in the original model development. The full dataset (N = 22) used for LOO-CV evaluation is available from the corresponding author upon request

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

The comparative analysis of activation functions and optimization algorithms (Table 4) revealed that the ANN architecture is highly sensitive to the choice of these hyperparameters. The hyperbolic tangent activation function (*tanh*), when combined with the Stochastic Gradient Descent (SGD) optimizer, yielded the lowest in-sample training error (MAPE = 3.93×10^{-6}; MSE = 2.28×10^{-10}), which reflect fitting to the training set rather than predictive generalization. Under Leave-One-Out Cross-Validation (LOO-CV, N = 22), the MLP tanh/SGD model yielded $R^2 = -0.119$, RMSE = 1705.4, and MAE = 1515.7 (mg AAE 100 g^{-1}), indicating that the (20-15-10) architecture does not generalize beyond the training data at this sample size. By contrast, Ridge Regression achieved R^2-LOO = 0.763, RMSE = 784.6, and MAE = 612.6, representing the best predictive performance among all evaluated models (Table 4).

Table 4. LOO-CV performance metrics (R^2, RMSE, MAE, MSE) for the MLP and benchmark models evaluated for antioxidant activity prediction (N = 22). In-sample MAPE values for MLP configurations are provided for descriptive comparison only and should not be interpreted as generalization metrics.

Model (Activation / Optimizer)	MAPE (%) in-sample	MSE in-sample	R^2 (LOO-CV)	RMSE (LOO-CV)	MAE (LOO-CV)	N
ReLU / SGD	3.04	19.09	−0.164	1747.2	1580.9	22
ReLU / Adam	0.0021	6.37×10^{-6}	−0.521	1969.3	1082.6	22
tanh / SGD	3.93×10^{-6}	2.28×10^{-10}	−0.119	1705.4	1515.7	22
tanh / Adam	1.10	8.93	−0.396	1904.8	1061.4	22
Logistic / SGD	0.08	0.03	−0.191	1771.3	1605.2	22
Logistic / Adam	2.33	47.41	−0.608	2024.7	1099.3	22
Ridge Regression	n/a	n/a	0.763	784.6	612.6	22
PLS (2 components)	n/a	n/a	0.726	844.0	610.1	22
PLS (3 components)	n/a	n/a	0.737	826.1	571.4	22

The predictive efficiency of the model can also be attributed to the careful selection of input variables. Compounds such as Gallic Acid and Quercetin presented higher concentrations in matrices such as jabuticaba and jambolão (Table 2). From a

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

biochemical standpoint, the neural network was able to learn that the presence of multiple hydroxyl groups in these molecules correlates directly with higher radical scavenging activity in the DPPH assay, a relationship well established in the structural-activity literature for polyphenolic compounds (PLATZER *et al.*, 2022; RUDRAPAL *et al.*, 2022). The retention of juçara in the full dataset (N = 22), despite its biochemically distinct profile, was a deliberate methodological choice: by including it as a held-out test observation in each LOO iteration, the framework quantifies prediction error transparently across all matrix types rather than artificially inflating performance by removing difficult samples. The resulting LOO-CV metrics therefore reflect true model behavior across the full compositional range of the dataset, including matrices where the 13 target markers underrepresent the actual antioxidant-relevant chemistry (GOODFELLOW; BENGIO; COURVILLE, 2017).

This outcome can be attributed to the zero-centered nature of the tanh function, which frequently accelerates gradient convergence compared to the logistic sigmoid function (DUBEY; SINGH; CHAUDHURI, 2022). In contrast, the *ReLU* function, although widely favored for its computational efficiency, yielded higher errors in this specific dataset (Figure 4), suggesting that the nonlinear relationship between phenolic compounds and DPPH response is better captured by smooth, saturating functions such as *tanh* (DUBEY; SINGH; CHAUDHURI, 2022).

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

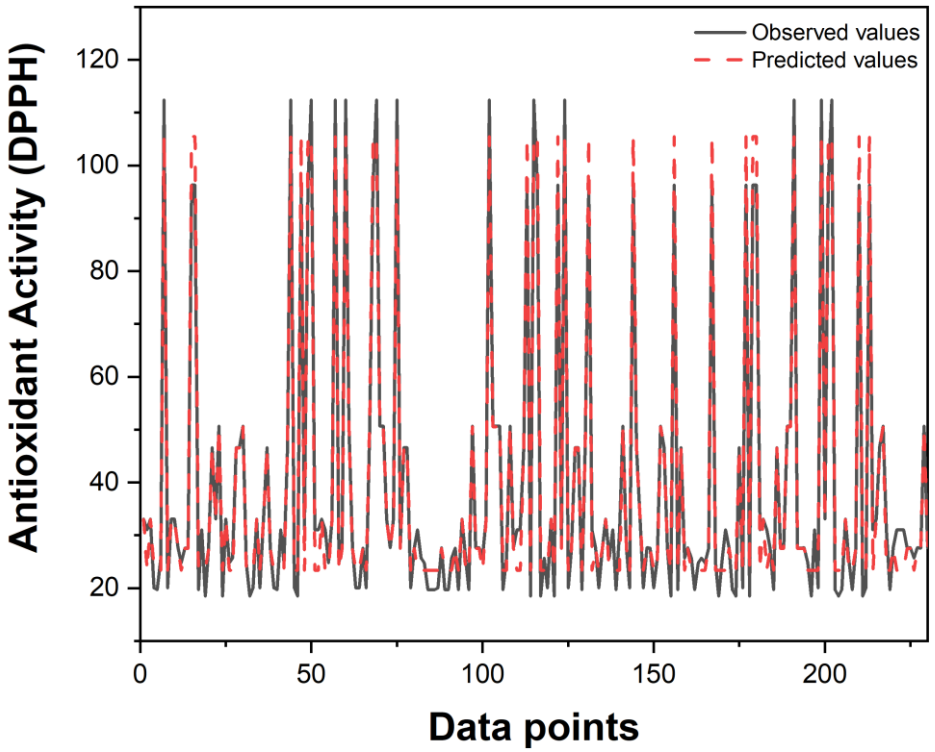

Figure 4. Comparison between observed and predicted antioxidant activity values (DPPH) obtained by the model with *ReLU* activation function and SGD optimizer, showing greater dispersion relative to the optimized model (MAPE = 3.04%; MSE = 19.09%).

Figure 5 illustrates the in-sample fit of the optimized MLP model, where predicted values closely track observed values within the training set. It is important to note that this apparent precision reflects model memorization rather than generalization: under LOO-CV, the same MLP architecture yielded $R^2 = -0.119$, indicating that predictions on unseen samples were systematically less accurate than a simple mean baseline. This outcome is consistent with the known susceptibility of high-complexity networks to overfitting when the number of training samples is small relative to model parameters (GOODFELLOW; BENGIO; COURVILLE, 2017).

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

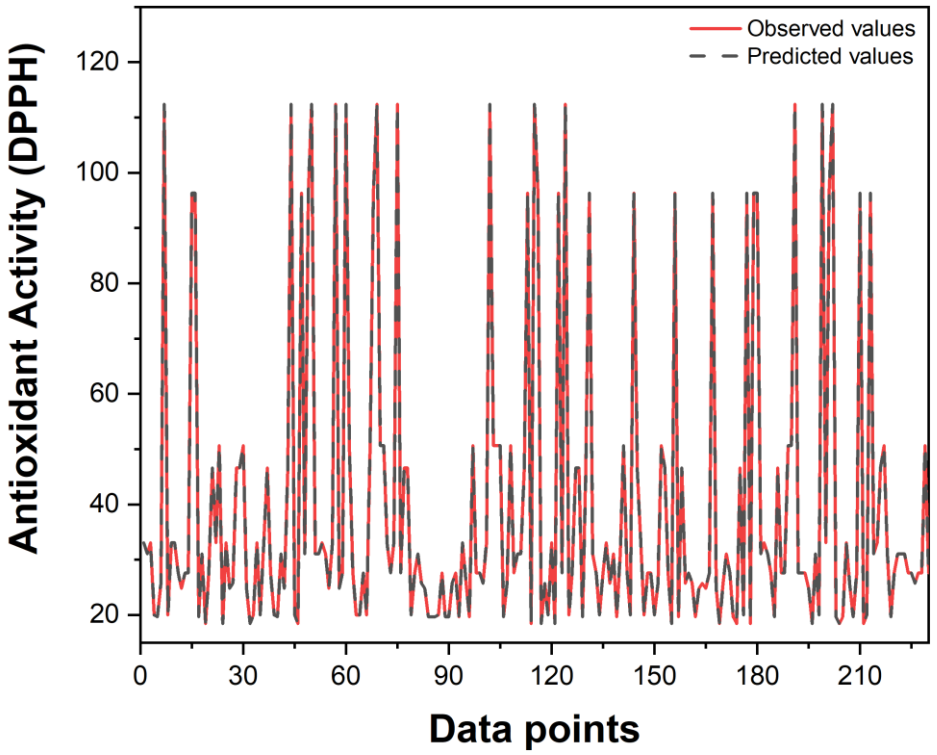

Figure 5. Comparison between observed and predicted antioxidant activity values (DPPH) obtained by the optimized model with hyperbolic tangent (*tanh*) activation function and SGD optimizer. In-sample fit: MAPE = 3.93×10^{-6}; MSE = 2.28×10^{-10} (training set only; LOO-CV: $R^2 = -0.119$, RMSE = 1705.4)

The LOO-CV results confirm that the small sample size (N = 22) is the central limiting factor in this study. With only N−1 = 21 training observations per LOO iteration, the MLP (20-15-10) topology, comprising hundreds of parameters, is severely underdetermined, leading to overfitting in which the network memorizes training patterns rather than capturing generalizable structure. The negative R^2-LOO values observed for both MLP configurations (tanh/SGD: −0.119; tanh/Adam: −0.396) quantify this limitation precisely: predictions on held-out samples were systematically worse than predicting the training mean. The benchmark analysis reveals that Ridge Regression, which regularizes parameter estimation and is well-suited to small high-dimensional datasets, achieves substantially better LOO-CV performance (R^2 = 0.763; RMSE = 784.6). This finding suggests that the relationship between phenolic profile and DPPH activity in this dataset is better captured by a linear regularized model than by the deep architecture evaluated, at least until a substantially larger

REMUNOM
ISSN 2178-6925
Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

prospective dataset becomes available. The present study should therefore be interpreted as an exploratory proof-of-concept investigation rather than a demonstration of MLP superiority for this chemometric problem.

Beyond antioxidant activity prediction, machine learning approaches hold broad potential in Food Engineering for applications such as thermal process optimization, sensory analysis, and quality screening. However, industrial deployment of ANN-based predictive tools requires prospective external validation with independently generated datasets, which remains an essential next step for future research.

5. Conclusion

The present study demonstrated the technical feasibility of predicting antioxidant capacity in diverse fruit matrices from phenolic profiles using machine learning. A comparative evaluation of MLP activation functions and optimizers was conducted, and the hyperbolic tangent (tanh) configuration with SGD produced the lowest in-sample training error. However, Leave-One-Out Cross-Validation (LOO-CV, N = 22) revealed that the MLP (20-15-10) architecture did not generalize beyond the training data (R^2-LOO = −0.119; RMSE = 1705.4 mg AAE 100 g^{-1}), consistent with the known sensitivity of high-complexity networks to overfitting in small-sample regimes. Ridge Regression emerged as the most defensible predictive model under current data constraints, achieving R^2-LOO = 0.763 and RMSE = 784.6, outperforming both MLP configurations and PLS regression across all LOO-CV metrics.

The full dataset (N = 22) was retained for model evaluation, including juçara samples whose phenolic profile is dominated by anthocyanins not quantified by the 13 target markers (Mann-Whitney U = 0; $p < 0.0001$ vs. remaining matrices for total phenolic content). Rather than excluding these samples, the LOO-CV framework treats them as independent test observations, allowing their prediction error to be reported transparently. The frequency-based variable selection criterion (≥ 22.7% occurrence across 22 matrices) provided a reproducible input screening protocol documented in full to support replication.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

The central limitation of this study is the small sample size (N = 22), which is insufficient to support the high-parameter complexity of MLP architectures and restricts generalization to other fruit matrices. Additional limitations include potential inter-study methodological heterogeneity in the compiled database and the absence of external prospective validation. For future work, it is recommended to: (i) expand the dataset with a greater diversity of fruit species and ripening stages using standardized analytical protocols; (ii) conduct prospective external validation with independently generated samples; (iii) apply interpretability methods such as permutation importance or SHAP values to identify the phenolic compounds most predictive of DPPH activity; and (iv) explore regularized architectures or ensemble methods designed for small-N chemometric datasets.

Acknowledgements

The authors thank the Coordination for the Improvement of Higher Education Personnel (CAPES), the National Council for Scientific and Technological Development (CNPq) (Processes 175901/2023-6 and 151329/2024-9).

References

AIDI WANNES, Wissem; MARZOUK, Brahim. Differences between Myrtle Fruit Parts (*Myrtus communis* var. *italica*) in Phenolics and Antioxidant Contents. *Journal of Food Biochemistry*, v. 37, n. 5, p. 585–594, out. 2013.

ALZUBAIDI, Laith *et al.* Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, v. 8, n. 1, p. 53, 31 mar. 2021.

BAI, Xuelian; ZHANG, Huawei; REN, Shuang. Antioxidant activity and HPLC analysis of polyphenol-enriched extracts from industrial apple pomace. *Journal of the Science of Food and Agriculture*, v. 93, n. 10, p. 2502–2506, 15 ago. 2013.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

BARROS DORES, Bruno Arley *et al.* GRAIN SIZE INFLUENCE ON THE EXTRACTION OF PHENOLIC COMPOUNDS AND ANTIMICROBIAL ACTIVITY IN TREE BARK FROM THREE BRAZILIAN BIOMES. *REMUNOM*, v. 13, n. 03, p. 1–26, 21 mar. 2026.

BETTA, Fabiana Della *et al.* Phenolic Compounds Determined by LC-MS/MS and In Vitro Antioxidant Capacity of Brazilian Fruits in Two Edible Ripening Stages. *Plant Foods for Human Nutrition*, v. 73, n. 4, p. 302–307, 14 dez. 2018.

BHAGYA RAJ, G. V. S.; DASH, Kshirod K. Comprehensive study on applications of artificial neural network in food process modeling. *Critical Reviews in Food Science and Nutrition*, v. 62, n. 10, p. 2756–2783, 5 abr. 2022.

CHANDIMALI, Nisansala *et al.* Free radicals and their impact on health and antioxidant defenses: a review. *Cell Death Discovery*, v. 11, n. 1, p. 19, 24 jan. 2025.

DEBRI, Rita Paola *et al.* Functional Foods as Vehicles for Bioactive Compounds: Chemical and Nutritional Perspectives on Health and Disease Prevention. *International Journal of Molecular Sciences*, v. 27, n. 5, p. 2293, 28 fev. 2026.

DIXIT, Versha *et al.* Functional Foods: Exploring the Health Benefits of Bioactive Compounds from Plant and Animal Sources. *Journal of Food Quality*, v. 2023, p. 1–22, 16 ago. 2023.

DUBEY, Shiv Ram; SINGH, Satish Kumar; CHAUDHURI, Bidyut Baran. Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, v. 503, p. 92–108, set. 2022.

FARIAS, Thayane Rabelo Braga; SANCHES, Natalia Beck; PETRUS, Rodrigo Rodrigues. The amazing native Brazilian fruits. *Critical Reviews in Food Science and Nutrition*, v. 26, n. 64, p. 9382–9399, 17 maio 2023.

FENG, Cheng-Yong *et al.* Polyphenol profile and antioxidant activity of the fruit and leaf of Vaccinium glaucoalbum from the Tibetan Himalayas. *Food Chemistry*, v. 219, p. 490–495, mar. 2017.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

GOODFELLOW, Ian.; BENGIO, Yoshua.; COURVILLE, Aaron. *Deep learning*. [S.l.]: The MIT Press, 2017.

GORBACHEV, Victor *et al.* Artificial Neural Networks for Predicting Food Antiradical Potential. *Applied Sciences*, v. 12, n. 12, p. 6290, 20 jun. 2022.

JUNG, Jinwoo *et al.* Predicting antioxidant activity of compounds based on chemical structure using machine learning methods. *The Korean Journal of Physiology & Pharmacology*, v. 28, n. 6, p. 527–537, 1 nov. 2024.

LIAKOS, Konstantinos G. *et al.* Machine Learning for Quality Control in the Food Industry: A Review. *Foods*, v. 14, n. 19, p. 3424, 4 out. 2025.

LOUZADA, Maria Laura da Costa *et al.* Consumo de alimentos ultraprocessados no Brasil: distribuição e evolução temporal 2008–2018. *Revista de Saúde Pública*, v. 57, n. 1, p. 12, 15 mar. 2023.

LV, Qi-zhuang *et al.* Current State of Knowledge on the Antioxidant Effects and Mechanisms of Action of Polyphenolic Compounds. *Natural Product Communications*, v. 16, n. 7, 31 jul. 2021.

MADHIARASAN, Manogaran; LOUZAZNI, Mohamed. Analysis of Artificial Neural Network: Architecture, Types, and Forecasting Applications. *Journal of Electrical and Computer Engineering*, v. 2022, p. 1–23, 18 abr. 2022.

NASCIMENTO, Eidla M.M. *et al.* HPLC and in vitro evaluation of antioxidant properties of fruit from Malpighia glabra (Malpighiaceae) at different stages of maturation. *Food and Chemical Toxicology*, v. 119, p. 457–463, set. 2018.

PARIKH, Bhumi; PATEL, V.H. Quantification of phenolic compounds and antioxidant capacity of an underutilized Indian fruit: Rayan [Manilkara hexandra (Roxb.) Dubard]. *Food Science and Human Wellness*, v. 6, n. 1, p. 10–19, mar. 2017.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

PEREIRA GONZATTO, Mateus; SCHERER SANTOS, Júlia. Introductory Chapter: Worldwide Fruit Crops Production and Research. In: GONZATTO, MATEUS PEREIRA; SANTOS, JÚLIA SCHERER (Org.). . *Fruit Crops Science - Ecophysiological and Horticultural Perspectives*. 1. ed. Internet: Intechopen, 2025. v. 1. .

PETRY, Henrique Belmonte *et al.* Brazilian Fruit Agribusiness. *Fruit Crops Science Journal*, v. 1, 2025.

PLATZER, Melanie *et al.* Radical Scavenging Mechanisms of Phenolic Compounds: A Quantitative Structure-Property Relationship (QSPR) Study. *Frontiers in Nutrition*, v. 9, 4 abr. 2022.

ROCCHETTI, Gabriele *et al.* Functional implications of bound phenolic compounds and phenolics–food interaction: A review. *Comprehensive Reviews in Food Science and Food Safety*, v. 21, n. 2, p. 811–842, 12 mar. 2022.

RUBERTO, Giuseppe; BARATTA, Maria T. Antioxidant activity of selected essential oil components in two lipid model systems. *Food Chemistry*, v. 69, n. 2, p. 167–174, maio 2000.

RUDRAPAL, Mithun *et al.* Dietary Polyphenols and Their Role in Oxidative Stress-Induced Human Diseases: Insights Into Protective Effects, Antioxidant Potentials and Mechanism(s) of Action. *Frontiers in Pharmacology*, v. 13, 14 fev. 2022.

SADOWSKA-BARTOSZ, Izabela; BARTOSZ, Grzegorz. Evaluation of The Antioxidant Capacity of Food Products: Methods, Applications and Limitations. *Processes*, v. 10, n. 10, p. 2031, 8 out. 2022.

SCHULZ, Mayara *et al.* Chemical composition, bioactive compounds and antioxidant capacity of juçara fruit (Euterpe edulis Martius) during ripening. *Food Research International*, v. 77, p. 125–131, nov. 2015.

REMUNOM
ISSN 2178-6925

Received: 10/03/2026 - **Accepted:** 15/04/2026
Vol: 13.06
DOI: 10.66104/z2na3j02
Pages: 1-30

SCHULZ, Mayara *et al.* Determination of Phenolic Compounds in Three Edible Ripening Stages of Yellow Guava (Psidium cattleianum Sabine) after Acidic Hydrolysis by LC-MS/MS. *Plant Foods for Human Nutrition*, v. 75, n. 1, p. 110–115, 7 mar. 2020.

SERAGLIO, Siluana Katia Tischer *et al.* Nutritional and bioactive potential of Myrtaceae fruits during ripening. *Food Chemistry*, v. 239, p. 649–656, jan. 2018.

SUN, Wenli; SHAHRAJABIAN, Mohamad Hesam. Therapeutic Potential of Phenolic Compounds in Medicinal Plants—Natural Health Products for Human Health. *Molecules*, v. 28, n. 4, p. 1845, 15 fev. 2023.

VEIGA, Izabella Paula Araújo *et al.* Fruit and vegetable consumption among Brazilian adults: trends from 2008 to 2023. *Cadernos de Saúde Pública*, v. 41, n. 1, 2025.

WANG, Yutang *et al.* Bioactive compounds and *in vitro* antioxidant activities of peel, flesh and seed powder of kiwi fruit. *International Journal of Food Science & Technology*, v. 53, n. 9, p. 2239–2245, 6 set. 2018.

ZHANG, Yuanyuan *et al.* A Brief Review of Phenolic Compounds Identified from Plants: Their Extraction, Analysis, and Biological Activity. *Natural Product Communications*, v. 17, n. 1, 5 jan. 2022.