

Article

Generalized ANN Model for Predicting the Energy Potential of Heterogeneous Waste

Ivan Brandić ¹, Ana Matin ^{1,*}, Karlo Špelić ¹, Nives Jovičić ², Božidar Matin ³, Mateja Grubor ¹ and Neven Voća ¹

¹ Faculty of Agriculture, University of Zagreb, Svetošimunska cesta 25, 10000 Zagreb, Croatia; ibrandic@agr.hr (I.B.); kspelic@agr.hr (K.Š.); mgrubor@agr.hr (M.G.); nvoca@agr.hr (N.V.)

² University of Applied Sciences Velika Gorica, 10410 Velika Gorica, Croatia; nives.jovicic@vvg.hr

³ Faculty of Forestry and Wood Technology, University of Zagreb, Svetošimunska cesta 23, 10000 Zagreb, Croatia; bmatin@sumfak.unizg.hr

* Correspondence: amatin@agr.hr

Abstract

In this paper, an artificial neural network (ANN) model of the MLP 5-17-1 type was developed to predict the gross calorific value (HHV) of various waste types based on ultimate analysis. The dataset comprised heterogeneous samples, including biomass, municipal and industrial waste, sludges, and derived fuels, ensuring the model's diversity and universality. The model achieved high accuracy ($R^2 = 0.92$; RMSE = 2.36; MAE = 1.68; MAPE = 10.99%), comparable to previous research results. The heterogeneity of the samples confirmed wide variations in composition and energy properties, which are crucial for developing a universal predictive model. The results confirm that ANN is a reliable tool for assessing the energy potential of waste and highlight the importance of expanding databases and optimizing parameters in future research.

Keywords: artificial neural network; data processing; energy potential; machine learning; waste



Academic Editor: Massimo Dentice D'Accadia

Received: 30 September 2025

Revised: 19 November 2025

Accepted: 20 November 2025

Published: 22 November 2025

Citation: Brandić, I.; Matin, A.; Špelić, K.; Jovičić, N.; Matin, B.; Grubor, M.; Voća, N. Generalized ANN Model for Predicting the Energy Potential of Heterogeneous Waste. *Energies* **2025**, *18*, 6111. <https://doi.org/10.3390/en18236111>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Waste management represents one of the biggest problems today [1], and it is directly related to the actions of people, whose number is rapidly increasing daily [2]. Although more technologies have been developed recently that try to reduce the mentioned problem, there is still no completely sustainable way of waste management [3]. In this context, the growing need for sustainable waste management and energy supply emphasizes the importance of the waste-to-energy process. The mentioned technologies enable the utilization of energy from different sources through thermochemical, biochemical, and chemical processes, whereby energy can be produced in different forms [4].

In this context, the principles of sustainable development of this concept consist of 5R (reduce, reuse, recycle, recovery, reclamation) and play a key role in the transition from the linear model of the economy based on the produce-use-discard pattern [5]. Emerging technologies for waste conversion and treatment have significant potential for producing clean energy and valuable products, reducing waste disposal, and supplementing recycling [6]. As waste is a valuable energy resource, it is necessary to know its characteristics, one of the most important being the higher heating value (HHV) [7].

HHV is determined in the laboratory using an adiabatic calorimeter [8]. However, it is important to note that this method requires personnel training, requires time, and is

often expensive. For this reason, various equations have been used to approximate HHV as accurately as possible [9]. To create such equations, data from ultimate analysis [10] are often used as input variables, consisting of five variables: carbon (C), hydrogen (H), nitrogen (N), oxygen (O), and sulfur (S) [11]. Due to the complexity of larger datasets and their interpretation, machine learning models have been increasingly applied in recent years for modeling, approximation, and predictive analysis [12]. One of the most widely used machine learning models is the artificial neural network (ANN), which operates on the principle of the human brain, making it suitable for analyzing complex tasks [13]. ANN models are suitable for modeling nonlinear tasks, identifying patterns between input and output variables, and are based on a layered structure through which data passes [14]. On the other hand, although suitable for modeling, ANN models require parameter tuning and optimization, which is not a simple task [15]. One of the most popular types of ANN models is the multilayer perceptron (MLP), a network in which information flows unidirectionally through the input, hidden, and output layers, which are interconnected [16]. A significant number of scientific papers focus on the development of ANN models for estimating the HHV of different materials based on ultimate analysis [17–21]. However, most research to date has focused on specific types of raw materials, such as biomass and coal, which can significantly limit the universality of the developed model, as this approach reduces the variability of input and output data. Previous research on predicting the higher calorific value (HHV) of biomass has revealed considerable differences in both the approaches and the quality of the data used. Abdollahi et al. [22] report that earlier studies randomly used datasets from ultimate and proximate analyses, as well as their combinations, resulting in heterogeneous and difficult-to-compare models. In contrast, Köcer [23] notes that newer machine learning algorithms have improved prediction accuracy, but most developed models are still based on limited datasets and specific types of biomass, which reduces their generalizability. In this research, a new methodology is presented that integrates a highly heterogeneous dataset collected from multiple independent sources and applies standardized preprocessing and an optimized artificial neural network (ANN) architecture to achieve a universal HHV prediction model, thereby overcoming the limitations of previous studies related to narrow datasets and limited model applicability. The aim of this research is to develop and validate an artificial neural network (ANN) for the reliable prediction of the higher heating value (HHV) of biomass from various origins, using elemental composition (C, H, N, S, O) as input variables. The research involved: (1) collection and standardization of data from published sources, (2) analysis of correlations between elements and HHV, (3) development, training, and evaluation of multiple ANN model architectures, and (4) comparison of ANN performance with traditional empirical models. The ultimate goal was to define a model with an optimal balance of accuracy and generalizability, applicable to a wide range of biomasses.

2. Materials and Methods

2.1. Data Collection

The data used to develop the model included various waste types, such as agricultural biomass residues, municipal waste, sludges, digestates, and treated waste, and were collected from published scientific papers [24–36]. Data on the ultimate analysis of the samples (percentages of C, H, N, S, and O), as well as the higher heating value (HHV), were collected. The samples cover a wide range of materials, including biomass of plant origin (wood, agricultural residues, food) and municipal and industrial waste (plastic, paper, textiles, rubber). Such diversity confirms the high heterogeneity of the samples used in model development. The raw data used for modeling can be seen in Supplementary Table S1.

2.2. Data Preprocessing

After the data were collected, a database was created. Due to the unevenness of the data, the next step was data cleaning, standardization, and normalization. Data cleaning is the process of identifying and removing errors from the database to increase quality and consistency [37]. The data were then standardized and normalized to prepare for the creation of artificial neural network models [38,39]. To fill data gaps based on known values, the min-max scaling method was used as preparation for K-Nearest Neighbor (kNN) methods [40,41]. After these methods were implemented, the data were ready for the creation of an ANN model.

2.3. Statistical Analysis

Statistical analysis was conducted on preprocessed data (after data cleaning). For this purpose, and for visualization, a pair plot diagram was created to display distributions and correlations. The characteristics of the research variables, i.e., the descriptive statistics metrics, were also presented numerically. Statistical analysis was performed using the Python 3.14 programming language [42].

2.4. Artificial Neural Networks—Model Settings

The ANN model is constructed with an architecture comprising input, hidden, and output layers [43]. The ultimate analysis variables were used as input data, and the higher heating value (HHV) was used as the output. The model was built as a Multilayer Perceptron (MLP) [44], as such models are most suitable for generalization [45]. The architecture, specifically the number of artificial neurons in the hidden layer, is determined randomly using the trial-and-error method [46–48]. The data was then divided into training, testing, and validation sets in a 70:15:15 ratio [49–51], which is the most commonly used. The model was created using the Python programming language [42] with associated packages [52–54]. The ANN was trained for 100,000 cycles to obtain relevant results.

Figure 1 shows the architecture of the developed ANN model, with five artificial neurons in the input layer, seventeen in the hidden layer, and one in the output layer.

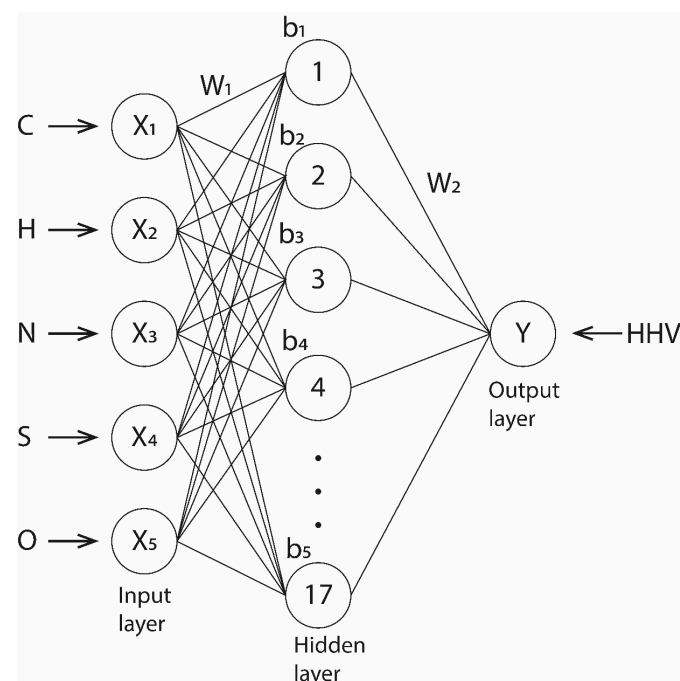


Figure 1. Architecture of the developed ANN model: MLP 5-17-1.

The output value of the model was calculated using the following Formula (1) [55]:

$$Y = f \cdot (b + W \cdot X) \quad (1)$$

where Y is the output value, f is the model function, b is the threshold, W is the weighting coefficient, and X is the model input value.

The selected methods were applied to achieve high accuracy and reliability in predicting the HHV of biomass. The artificial neural network was chosen for its ability to model complex nonlinear relationships between elemental composition (C, H, N, S, O) and HHV, which classical regression approaches cannot fully capture. Standardization, normalization, and imputation of data using the k-nearest neighbors' method were performed to reduce the impact of outliers and information loss. Dividing the dataset into training, validation, and test sets (70:15:15) enabled assessment of the model's generalization ability and prevented overfitting, thus ensuring the reliability and repeatability of the results.

2.5. Model Evaluation

After modeling to estimate the output value, the model needed to be evaluated. To assess the success of the model in terms of modeling error and specific regression indicators, the following metrics were used: root mean squared error (RMSE) (2), mean absolute error (MAE) (3) [56], coefficient of determination (R^2) (4), and mean absolute percentage error (MAPE) (5) [57]:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (3)$$

$$R^2 = \frac{\sum_{i=1}^n (X_i - Y_i)^2}{\sum_{i=1}^n (\bar{Y} - Y)^2} \quad (4)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - X_i}{Y_i} \right| \quad (5)$$

3. Results

Figure 2 presents a graphical representation of the scatterplot matrix and the mutual relationships among the observed variables used as input and output variables in the model.

Table 1 presents the basic statistical characteristics of the variables used as input and output data for the model. The statistical indicators include variability, mean value, and standard deviation.

Table 1 presents the descriptive statistics of the observed variables used to develop the ANN model.

Table 2 shows a summary of the developed ANN model 5-17-1.

Table 2 presents a summary of the developed ANN model, a multilayer perceptron with 5 input neurons, 17 hidden neurons, and 1 output neuron per layer, and displays the performance and model errors.

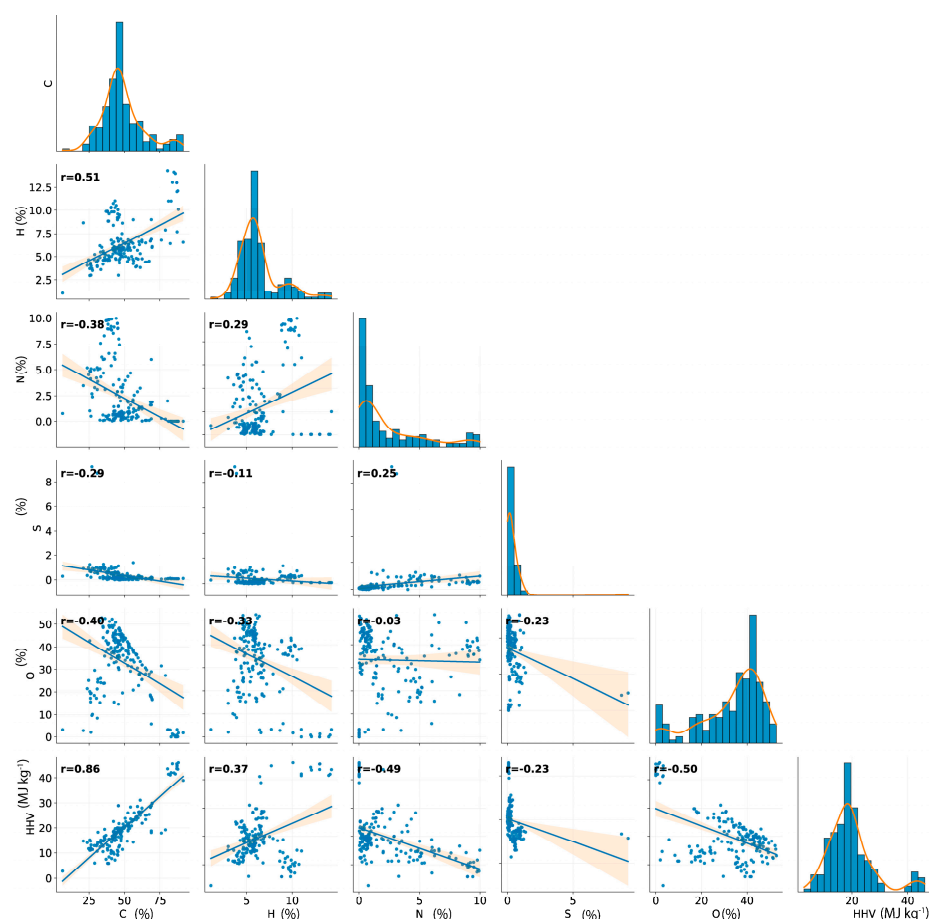


Figure 2. Scatterplot matrix with pairwise relationships and correlations between the input variables and the output (HHV).

Table 1. Descriptive statistics of the observed dataset after the cleaning, standardization, and normalization process.

Descriptive Statistics ↓	Input Variables (%) ↓					Output Variable (MJ kg ⁻¹) ↓
	C (%)	H (%)	N (%)	S (%)	O (%)	HHV
Mean	48.62	6.37	2.35	0.37	33.35	19.59
Median	45.78	5.90	0.90	0.13	37.94	18.75
Minimum	6.27	1.09	0.01	0.00	0.00	2.80
Maximum	91.53	14.30	9.98	9.20	52.84	46.08
Std.Dev.	14.76	2.24	2.83	0.91	13.66	8.35

C—content of carbon; H—content of hydrogen; N—content of nitrogen; S—content of sulfur; O—content of oxygen; HHV—higher heating value.

Table 2. Summary of the developed ANN MLP 5-17-1 model.

ANN Model	Performance (R ²)			Error		
	Training	Test	Validation	Training	Test	Validation
MLP 5-17-1	0.94	0.85	0.86	2.06	4.96	3.99

ANN—Artificial neural networks; MLP—Multilayer perceptron; R²—Coefficient of determination.

Table 3 presents the model reliability metrics, including model error (RMSE, MAE, and MAPE) and the specific regression indicator (R²).

Table 3. Performance evaluation of the ANN model based on training, validation, test, and overall datasets.

Dataset	R ²	RMSE	MAE	MAPE (%)
Train	0.943	2.03	1.49	8.96
Validation	0.856	2.82	1.91	15.10
Test	0.85	3.15	2.30	16.48
Overall	0.92	2.36	1.68	10.99

R²—Coefficient of determination; RMSE—Root mean squared error; MAE—Mean absolute error; MAPE—Mean absolute percentage error.

Figure 3 shows a comparison of the predicted and actual HHV values for the training, validation, and testing datasets. The histograms along the axes display the distribution of the target and predicted values.

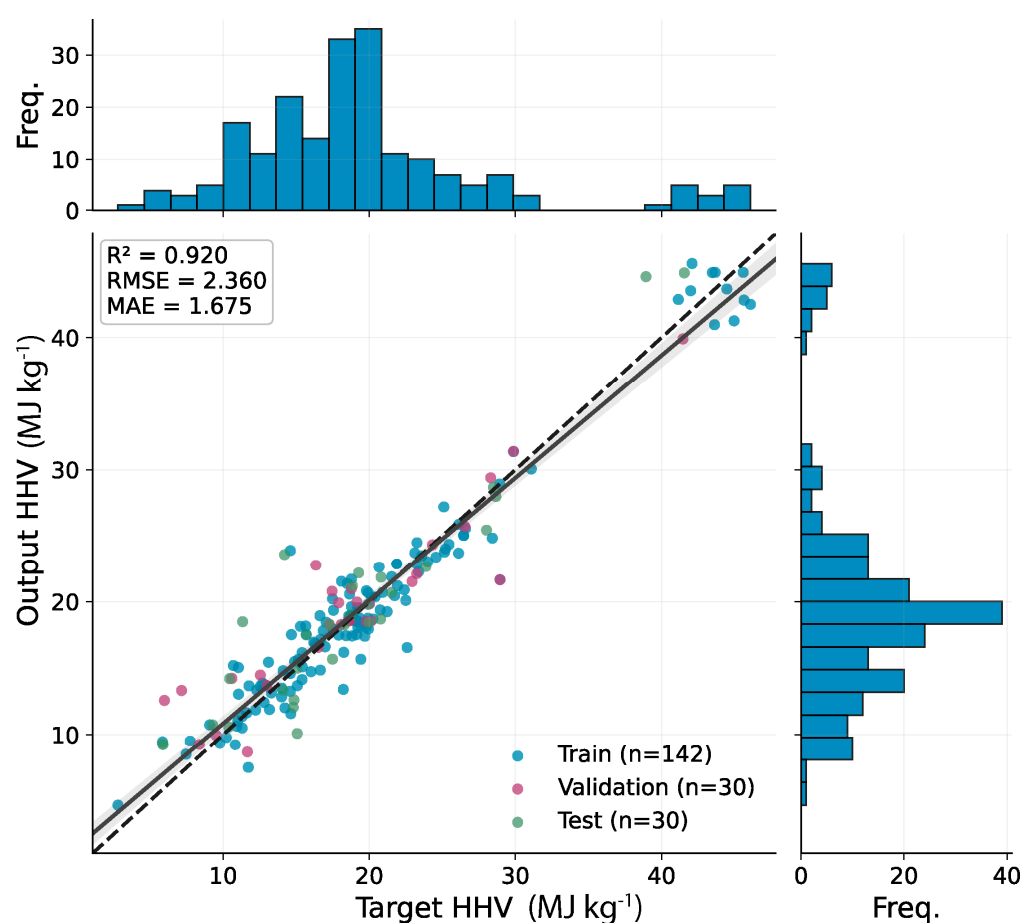
**Figure 3.** Comparison of predicted and observed HHV values with marginal distributions.

Figure 4 shows the relative importance of the input variables of the ultimate analysis in the ANN model for HHV prediction. Relative importance refers to dimensionless values obtained from the ANN model sensitivity analysis, in which each variable is assigned a numerical importance index indicating its relative contribution to the model's predictive accuracy.

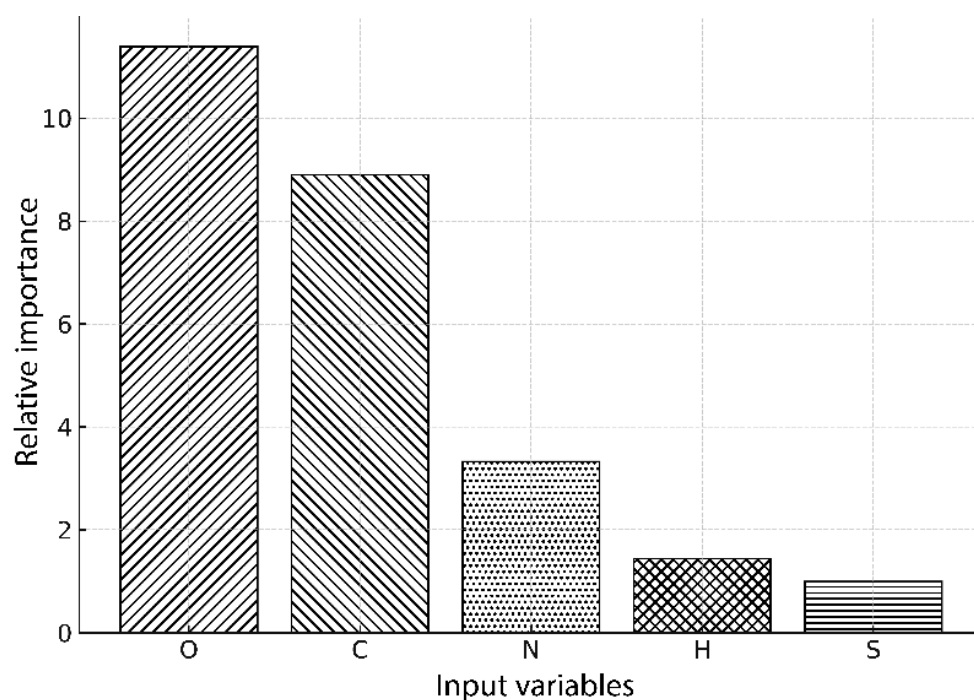


Figure 4. Relative importance of input variables in the ANN model (dimensionless sensitivity values).

4. Discussion

Scatterplot matrix analysis (Figure 1) shows that carbon content has a strong positive correlation with HHV ($r = 0.86$). García-Nieto et al. [58] conducted a correlation analysis of ultimate analysis and HHV of biomass in their research, and also found a strong positive correlation with carbon content. High carbon content indicates that samples have higher HHV [59]. Table 1 shows that the average carbon content in the samples analyzed is 48.62%, while the average HHV is 19.59 MJ kg^{−1}. The range of values indicates considerable heterogeneity in the data—for example, carbon content varies from 6.27% to 91.53%, and HHV from 2.80 to 46.08 MJ/kg. Esteves et al. [60] report that different types of biomass potentially usable as fuels exhibit substantial variation in their physicochemical characteristics and provide HHV values for wood in the range of 17–23 MJ kg^{−1}. Large variations in the proportions of individual elements and HHV result from the heterogeneity of the samples [61,62], as they are different types of waste, which was crucial for creating a more universal model with high generalization ability. For these reasons, ANN models, which are machine learning algorithms inspired by the structure and function of the human brain, are considered the most suitable for solving nonlinear problems involving large amounts of nonlinear data [63]. The ANN model developed in this research achieved coefficients of determination (R^2) of 0.942 for the training dataset, 0.86 for the validation dataset, and 0.85 for the test dataset, indicating good generalization ability. The error values (2.069 for training, 3.990 for validation, and 4.964 for test) confirm the stability of the model without signs of overfitting.

The best results are achieved by the model on the training dataset ($R^2 = 0.943$; RMSE = 2.03), while slightly lower values are obtained on the validation ($R^2 = 0.856$) and test datasets ($R^2 = 0.85$), indicating good, but not perfect, generalization. Overall, the model achieves high accuracy ($R^2 = 0.92$) with acceptable error values (RMSE = 2.36; MAE = 1.68; MAPE = 10.99%), confirming its reliability in predicting the HHV of different waste types. The overall evaluation of the model shows high accuracy with $R^2 = 0.92$, indicating a strong correlation between the predictions and the actual HHV values. The error values (RMSE = 2.36; MAE = 1.68; MAPE = 10.99%) confirm that the model reliably

predicts the energy potential of different waste types, with an acceptable level of deviation. Olatunji et al. [64] report RMSE of 3.587 and MAPE of 21.68. Dashti et al. [65] conducted a study comparing several machine learning models for estimating the higher heating value (HHV) of biomass using input data from proximate analysis. The ANN achieved a high regression indicator (R^2) for training (0.94) and testing (0.95), and a low total error MSE (0.83), indicating the suitability of this model for this problem. In the research by Mondal and Rafizul [66], the ANN model was used to predict the calorific value of municipal waste based on proximate analysis (MC, VC, FC, Ash). After 1000 epochs, the model achieved high accuracy with $R^2 = 0.9397$, $MSE = 1.599$, $RMSE = 1.264$, $MAD = 0.558$, and $MAPE = 0.032$, confirming its ability to accurately model nonlinear relationships between the chemical composition of waste and its energy value.

Matveeva and Byckov [67] conducted research to create an ANN model for estimating the HHV of biofuels and obtained R^2 values of 0.82–0.87, considering the data management process. Veza et al. [68] report R^2 values for their model of 0.80–0.94. From all the above, it is evident that the created ANN MLP 5-17-1 model showed satisfactory performance in predicting HHV. The greatest influence on the model is exerted by the proportion of oxygen (O), followed by carbon (C), while nitrogen (N), hydrogen (H), and sulfur (S) are significantly less important. The results indicate that oxygen and carbon are the key predictors of the energy value of different types of waste, which is consistent with established patterns in the literature. Brandić et al. [18] state that the output value of HHV is mostly influenced by an increase in the variable S, while Adeleke et al. [69] report the highest influence from ash, sulfur, nitrogen, and oxygen. However, the varying influence of individual variables is determined by differences in samples, i.e., heterogeneity, as different types of waste are characterized by specific chemical compositions and energy properties [69–72]. These results are consistent with previous studies that demonstrated the effectiveness of ANN in modeling HHV based on proximate or ultimate analysis data. Compared to previous research, the developed model demonstrates equal or higher accuracy and stability, with improved generalization ability due to the broader range and greater heterogeneity of the data used.

Although the model demonstrates high performance and a high level of universality or generalization, there are key limitations that must be considered. As previously mentioned, sample heterogeneity leads to variability in the chemical composition of the analyzed samples, which can reduce model accuracy, particularly in certain subgroups. Therefore, planning and optimization are essential [73]. It is also important to emphasize that model performance depends on the quality and resolution of the input variables, as models are sensitive to poor-quality data [74]. In future research, it is advisable to build models using different data inputs while increasing the quality and quantity of the entire dataset [75]. Another key factor that can improve problem-solving is hyperparameter optimization and regularization, which can further reduce prediction errors [76].

5. Conclusions

In this study, an artificial neural network (ANN) model was developed and validated to predict the higher heating value (HHV) of various types of biomass and waste based on their elemental composition (C, H, N, S, O). The model demonstrated high accuracy ($R^2 = 0.92$; $RMSE = 2.36$ MJ/kg), with stable performance across training, validation, and test sets, confirming its reliability and generalizability. Analysis of variable importance showed that oxygen (O) and carbon (C) have the greatest impact on HHV, consistent with established energy laws of biomass. The scientific novelty of this work lies in the integration of a highly heterogeneous dataset collected from multiple sources and the application of standardized preprocessing and optimized ANN model architecture, enabling universal prediction of

HHV for different types of biomass. The practical significance of the results is evident in the ability to rapidly assess the energy potential of waste and biomass without complex laboratory analyses, contributing to more efficient planning of energy production from renewable sources and the development of a circular economy.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/en18236111/s1>. Table S1: Raw data for modeling.

Author Contributions: Conceptualization, I.B. and A.M.; methodology, K.Š.; software, I.B.; validation, N.V., A.M., and K.Š.; formal analysis, M.G.; investigation, N.J.; data curation, B.M.; writing—original draft preparation, I.B.; writing—review and editing, A.M.; visualization, I.B.; supervision, N.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Acknowledgments: During the preparation of this manuscript, the author(s) used software tools such as InstaText, Grammarly, and QuillBot exclusively to improve sentence syntax, readability, and grammatical correctness. These tools were not used to generate figures, data, or substantive content.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Gemar, G.; Soler, I.P.; Sánchez-Teba, E.M. Waste Management: Valorisation Is the Way. *Foods* **2021**, *10*, 2373. [\[CrossRef\]](#)
- Czekala, W.; Drozdowski, J.; Łabiak, P. Modern Technologies for Waste Management: A Review. *Appl. Sci.* **2023**, *13*, 8847. [\[CrossRef\]](#)
- Kumari, T.; Raghubanshi, A.S. Waste Management Practices in the Developing Nations: Challenges and Opportunities. In *Waste Management and Resource Recycling in the Developing World*; Elsevier: Amsterdam, The Netherlands, 2023; pp. 773–797.
- Lisbona, P.; Pascual, S.; Pérez, V. Waste to Energy: Trends and Perspectives. *Chem. Eng. J. Adv.* **2023**, *14*, 100494. [\[CrossRef\]](#)
- Sharma, S.; Basu, S.; Shetti, N.P.; Kamali, M.; Walvekar, P.; Aminabhavi, T.M. Waste-to-Energy Nexus: A Sustainable Development. *Environ. Pollut.* **2020**, *267*, 115501. [\[CrossRef\]](#)
- Adeleke, O.; Akinlabi, S.A.; Jen, T.-C.; Dunmade, I. Sustainable Utilization of Energy from Waste: A Review of Potentials and Challenges of Waste-to-Energy in South Africa. *Int. J. Green Energy* **2021**, *18*, 1550–1564. [\[CrossRef\]](#)
- Kalivodová, M.; Baláš, M.; Milčák, P.; Lisá, H.; Lisý, M.; Lachman, J.; Kracík, P.; Křižan, P.; Vejražka, K. Determination of Higher Heating Value by Calculation Based on Elemental Analysis. *Paliva* **2022**, *14*, 8–20. [\[CrossRef\]](#)
- Acar, S.; Ayanoglu, A. Determination of Higher Heating Values (HHVs) of Biomass Fuels. *Energy Educ. Sci. Technol. Part A Energy Sci. Res.* **2012**, *28*, 749–758.
- Kwaghger, A.; Iortyer, L.A. The Development of Equations for Estimating High Heating Values from Proximate and Ultimate Analysis for Some Selected Indigenous Fuel Woods. *Eur. J. Eng. Technol.* **2017**, *5*, 21–33.
- Aghel, B.; Yahya, S.I.; Rezaei, A.; Alobaid, F. A Dynamic Recurrent Neural Network for Predicting Higher Heating Value of Biomass. *Int. J. Mol. Sci.* **2023**, *24*, 5780. [\[CrossRef\]](#)
- Kannah, R.Y.; Merrylin, J.; Devi, T.P.; Kavitha, S.; Sivashanmugam, P.; Kumar, G.; Banu, J.R. Food Waste Valorization: Biofuels and Value Added Product Recovery. *Bioresour. Technol. Rep.* **2020**, *11*, 100524. [\[CrossRef\]](#)
- Tyralis, H.; Papacharalampous, G.; Langousis, A. A Brief Review of Random Forests for Water Scientists and Practitioners and Their Recent History in Water Resources. *Water* **2019**, *11*, 910. [\[CrossRef\]](#)
- Gómez-Ramos, E.; Venegas-Martínez, F. Analítica: A Review of Artificial Neural Networks—How Well Do They Perform in Forecasting Time Series? *Analítica Rev. Análisis Estadístico* **2013**, *6*, 7–18.
- Kononenko, I. Machine Learning for Medical Diagnosis: History, State of the Art and Perspective. *Artif. Intell. Med.* **2001**, *23*, 89–109. [\[CrossRef\]](#)
- Gupta, T.K.; Raza, K. Optimization of ANN Architecture: A Review on Nature-Inspired Techniques. In *Machine Learning in Bio-Signal Analysis and Diagnostic Imaging*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 159–182. [\[CrossRef\]](#)
- Muñoz-Zavala, A.E.; Macías-Díaz, J.E.; Alba-Cuéllar, D.; Guerrero-Díaz-De-León, J.A. A Literature Review on Some Trends in Artificial Neural Networks for Modeling and Simulation with Time Series. *Algorithms* **2024**, *17*, 76. [\[CrossRef\]](#)
- Afolabi, I.C.; Epelle, E.I.; Gunes, B.; Güleç, F.; Okolie, J.A. Data-Driven Machine Learning Approach for Predicting the Higher Heating Value of Different Biomass Classes. *Clean Technol.* **2022**, *4*, 1227–1241. [\[CrossRef\]](#)

18. Brandić, I.; Pezo, L.; Bilandžija, N.; Peter, A.; Šurić, J.; Voća, N. Artificial Neural Network as a Tool for Estimation of the Higher Heating Value of Miscanthus Based on Ultimate Analysis. *Mathematics* **2022**, *10*, 3732. [\[CrossRef\]](#)
19. Ighalo, J.O.; Igwegbe, C.A.; Adeniyi, A.G. Multi-Layer Perceptron Artificial Neural Network (MLP-ANN) Prediction of Biomass Higher Heating Value (HHV) Using Combined Biomass Proximate and Ultimate Analysis Data. *Model. Earth Syst. Environ.* **2021**, *8*, 3177–3191. [\[CrossRef\]](#)
20. Zhu, X.; Yang, G. Study on HHV Prediction of Municipal Solid Wastes: A Machine Learning Approach. *Int. J. Energy Res.* **2021**, *46*, 3663–3673. [\[CrossRef\]](#)
21. Mondal, C.; Pal, S.K.; Samanta, B.; Dutta, D.; Raj, S. Analysis and Significance of Prediction Models for Higher Heating Value of Coal: An Updated Review. *J. Therm. Anal. Calorim.* **2023**, *148*, 7521–7538. [\[CrossRef\]](#)
22. Abdollahi, S.A.; Ranjbar, S.F.; Jahromi, D.R. Applying Feature Selection and Machine Learning Techniques to Estimate the Biomass Higher Heating Value. *Sci. Rep.* **2023**, *13*, 16093. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Kocer, A. Prediction of the Higher Heating Values of Biomass Using Machine Learning Methods Based on Proximate and Ultimate Analysis. *J. Mech. Sci. Technol.* **2024**, *38*, 1569–1574. [\[CrossRef\]](#)
24. Das, B.K.; Kader, M.A.; Hoque, S.M.N. Energy Recovery Potential from Municipal Solid Waste in Rajshahi City by Landfill Technique. *Int. J. Renew. Energy Res.* **2014**, *4*, 349–354.
25. Elwan, A.; Arief, Y.Z.; Adzis, Z.; Muhamad, N.A. Life Cycle Assessment-Based Environmental Impact Comparative Analysis of Composting and Electricity Generation from Solid Waste. *Energy Procedia* **2015**, *68*, 186–194. [\[CrossRef\]](#)
26. Pasek, A.D.; Gultom, K.W.; Suwono, A. Feasibility of Recovering Energy from Municipal Solid Waste to Generate Electricity. *J. Eng. Technol. Sci.* **2013**, *45*, 241–256. [\[CrossRef\]](#)
27. Fiori, L.; Basso, D.; Castello, D.; Baratieri, M. Hydrothermal Carbonization of Biomass: Design of a Batch Reactor and Preliminary Experimental Results. *Chem. Eng. Trans.* **2014**, *37*, 55–60. [\[CrossRef\]](#)
28. Yang, X.; Li, H.; Wang, Y.; Qu, L. Predicting Higher Heating Value of Sewage Sludges via Artificial Neural Network Based on Proximate and Ultimate Analyses. *Water* **2023**, *15*, 674. [\[CrossRef\]](#)
29. Samad, N.A.F.A.; Jamin, N.A.; Saleh, S. Torrefaction of Municipal Solid Waste in Malaysia. *Energy Procedia* **2017**, *138*, 313–318. [\[CrossRef\]](#)
30. Gañan, J.; Abdulla, A.A.-K.; Miranda, A.B.; Turegano, J.; Correia, S.; Cuerda, E.M. Energy Production by Means of Gasification Process of Residuals Sourced in Extremadura (Spain). *Renew. Energy* **2005**, *30*, 1759–1769. [\[CrossRef\]](#)
31. Cahyanti, M.N.; Doddapaneni, T.R.K.C.; Madissoo, M.; Pärn, L.; Virro, I.; Kikas, T. Torrefaction of Agricultural and Wood Waste: Comparative Analysis of Selected Fuel Characteristics. *Energies* **2021**, *14*, 2774. [\[CrossRef\]](#)
32. Hu, J.; Yu, F.; Lu, Y. Application of Fischer–Tropsch Synthesis in Biomass to Liquid Conversion. *Catalysts* **2012**, *2*, 303–326. [\[CrossRef\]](#)
33. Gerner, G.; Meyer, L.; Wanner, R.; Keller, T.; Krebs, R. Sewage Sludge Treatment by Hydrothermal Carbonization: Feasibility Study for Sustainable Nutrient Recovery and Fuel Production. *Energies* **2021**, *14*, 2697. [\[CrossRef\]](#)
34. Enebe, N.L.; Chigor, C.B.; Obileke, K.; Lawal, M.S.; Enebe, M.C. Biogas and Syngas Production from Sewage Sludge: A Sustainable Source of Energy Generation. *Methane* **2023**, *2*, 192–217. [\[CrossRef\]](#)
35. Zalazar-Garcia, D.; Fernandez, A.; Rodriguez-Ortiz, L.; Rodriguez, R.; Mazza, G. Sustainable Slow-Pyrolysis Simulation of 12 Lignocellulosic Bio-Wastes: CO₂ Emission, Energy, and Water Consumption. *IOP Conf. Ser. Earth Environ. Sci.* **2022**, *952*, 012008. [\[CrossRef\]](#)
36. Nhuchhen, D.; Afzal, M. HHV Predicting Correlations for Torrefied Biomass Using Proximate and Ultimate Analyses. *Bioengineering* **2017**, *4*, 7. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Wangikar, V.C.; Deshmukh, R.R. Data Cleaning: Current Approaches and Issues. *IEEE Data Eng. Bull.* **2015**, *23*, 3–13.
38. Jamal, P.; Ali, M.; Faraj, R.H.; Ali, P.J.M. Data Normalization and Standardization: A Technical Report. *Mach. Learn. Tech. Rep.* **2014**, *1*, 1–6.
39. Vinai, S. Standardization in Machine Learning. *Delhi Technol. Univ.* **2021**, *67*, 173–196.
40. Ali, P.J.M. Investigating the Impact of Min-Max Data Normalization on the Regression Performance of K-Nearest Neighbor with Different Similarity Measurements. *ARO Sci. J. Koya Univ.* **2022**, *10*, 85–91. [\[CrossRef\]](#)
41. Bhardwaj, C.A.; Mishra, M.; Desikan, K. Dynamic Feature Scaling for K-Nearest Neighbor Algorithm. *arXiv* **2022**. [\[CrossRef\]](#)
42. van Rossum, G.; Python Development Team. *Python Tutorial Release 3.7.0*, Python 3.7.0; Python Software Foundation: Beaverton, OR, USA, 2018; pp. 1–155.
43. Apicella, A.; Isgrò, F.; Prevete, R. Hidden Classification Layers: Enhancing Linear Separability between Classes in Neural Networks Layers. *Pattern Recognit. Lett.* **2023**, *177*, 69–74. [\[CrossRef\]](#)
44. Francik, S.; Łapczyńska-Kordon, B.; Pedryc, N.; Szewczyk, W.; Francik, R.; Ślipek, Z. The Use of Artificial Neural Networks for Determining Values of Selected Strength Parameters of Miscanthus × Giganteus. *Sustainability* **2022**, *14*, 3062. [\[CrossRef\]](#)
45. Agahian, S.; Akan, T. Battle Royale Optimizer for Training Multi-Layer Perceptron. *Evol. Syst.* **2021**, *13*, 563–575. [\[CrossRef\]](#)

46. Tufail, S.; Riggs, H.; Tariq, M.; Sarwat, A.I. Advancements and Challenges in Machine Learning: A Comprehensive Review of Models, Libraries, Applications, and Algorithms. *Electronics* **2023**, *12*, 1789. [CrossRef]
47. Hussein, M.M.; Mutlag, A.H.; Shareef, H. An Improved Artificial Neural Network Design for Face Recognition Utilizing Harmony Search Algorithm. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *745*, 012038. [CrossRef]
48. Bouzeria, H.; Ghenim, A.N.; Khanchoul, K. Using Artificial Neural Network (ANN) for Prediction of Sediment Loads, Application to the Mellah Catchment, Northeast Algeria. *J. Water Land Dev.* **2017**, *33*, 47–55. [CrossRef]
49. Kovačević, S.; Banjac, M.K.; Podunavac-Kuzmanović, S. Artificial Intelligence and Experimental Design: The Flywheel of Innovating Food Processing Engineering. *Processes* **2025**, *13*, 846. [CrossRef]
50. Jiang, Y. Prediction of Monthly Mean Daily Diffuse Solar Radiation Using Artificial Neural Networks and Comparison with Other Empirical Models. *Energy Policy* **2008**, *36*, 3833–3837. [CrossRef]
51. Chi, J.; Liu, Y.; Wang, V.; Yan, J. Performance Analysis of Three Kinds of Neural Networks in the Classification of Mask Images. *J. Phys. Conf. Ser.* **2022**, *2181*, 012032. [CrossRef]
52. Hunter, J.; Dale, D.; Firing, E.; Droettboom, M. Matplotlib: Release 2.0.2 John. 2017, 1–3148. Available online: <https://matplotlib.org/2.0.2/Matplotlib.pdf> (accessed on 1 June 2025).
53. NumPy Community. NumPy User Guide 1.22.4. 2022, 109. Available online: <https://numpy.org/doc/1.22/numpy-user.pdf> (accessed on 1 June 2025).
54. McKinney, W. Pandas: Powerful Python Data Analysis Toolkit. Pandas—Powerful Python Data Analysis Toolkit. 2022, 1–3743. Available online: <https://pandas.pydata.org/pandas-docs/version/0.7.3/pandas.pdf> (accessed on 1 June 2025).
55. Maqsoom, A.; Aslam, B.; Gul, M.E.; Ullah, F.; Kouzani, A.Z.; Mahmud, M.a.P.; Nawaz, A. Using Multivariate Regression and ANN Models to Predict Properties of Concrete Cured under Hot Weather. *Sustainability* **2021**, *13*, 10164. [CrossRef]
56. Chai, T.; Draxler, R.R. Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)?—Arguments against Avoiding RMSE in the Literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [CrossRef]
57. Chicco, D.; Warrens, M.J.; Jurman, G. The Coefficient of Determination R-Squared Is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [CrossRef]
58. García-Nieto, P.J.; García-Gonzalo, E.; Paredes-Sánchez, J.P. Estimation of the Coal Higher Heating Value for Energy Systems Relied on Ultimate Analysis with Machine Learning Techniques. *Fuel* **2023**, *357*, 130037. [CrossRef]
59. Sedai, P.; Kalita, D.; Deka, D. Assessment of the Fuel Wood of India: A Case Study Based on Fuel Characteristics of Some Indigenous Species of Arunachal Pradesh. *Energy Sources Part A Recover. Util. Environ. Eff.* **2016**, *38*, 891–897. [CrossRef]
60. Esteves, B.; Sen, U.; Pereira, H. Influence of Chemical Composition on Heating Value of Biomass: A Review and Bibliometric Analysis. *Energies* **2023**, *16*, 4226. [CrossRef]
61. Uler-Zefikj, M.; Godyń, K.; Tokarczyk, K.; Filkoski, R.V. Characterization of Municipal Solid Waste as Potential Fuel for Energy Needs. *Materials* **2025**, *18*, 2103. [CrossRef]
62. Boumanchar, I.; Chhiti, Y.; Alaoui, F.E.M.; Sahibed-Dine, A.; Bentiss, F.; Jama, C.; Bensitel, M. Municipal Solid Waste Higher Heating Value Prediction from Ultimate Analysis Using Multiple Regression and Genetic Programming Techniques. *Waste Manag. Res. J. A Sustain. Circ. Econ.* **2018**, *37*, 578–589. [CrossRef]
63. Kariri, E.; Louati, H.; Louati, A.; Masmoudi, F. Exploring the Advancements and Future Research Directions of Artificial Neural Networks: A Text Mining Approach. *Appl. Sci.* **2023**, *13*, 3186. [CrossRef]
64. Olatunji, O.O.; Akinlabi, S.; Madushele, N.; Adediji, P.A.; Felix, I. Multilayer Perceptron Artificial Neural Network for the Prediction of Heating Value of Municipal Solid Waste. *AIMS Energy* **2019**, *7*, 944–956. [CrossRef]
65. Dashti, A.; Noushabadi, A.S.; Raji, M.; Razmi, A.; Ceylan, S.; Mohammadi, A.H. Estimation of Biomass Higher Heating Value (HHV) Based on the Proximate Analysis: Smart Modeling and Correlation. *Fuel* **2019**, *257*, 115931. [CrossRef]
66. Mondal, S.; Rafizul, I.M. Predicting Calorific Value through Proximate Analysis of Municipal Solid Waste Using Soft Computing System. *Discov. Appl. Sci.* **2025**, *7*, 212. [CrossRef]
67. Matveeva, A.; Bychkov, A. How to Train an Artificial Neural Network to Predict Higher Heating Values of Biofuel. *Energies* **2022**, *15*, 7083. [CrossRef]
68. Veza, I.; Irianto, N.; Panchal, H.; Paristiawan, P.A.; Idris, M.; Fattah, I.M.R.; Putra, N.R.; Silambarasan, R. Improved Prediction Accuracy of Biomass Heating Value Using Proximate Analysis with Various ANN Training Algorithms. *Results Eng.* **2022**, *16*, 100688. [CrossRef]
69. Adeleke, O.; Olatunji, O.O.; Jen, T.-C.; Olawuyi, I. Enhanced Prediction of Heating Value of Municipal Solid Waste Using Hybrid Neuro-Fuzzy Model and Decision Tree-Based Feature Importance Assessment. *Green Energy Resour.* **2025**, *3*, 100119. [CrossRef]
70. Baziar, M.; Yousefi, M.; Oskoei, V.; Makhdoomi, A.; Abdollahzadeh, R.; Dehghan, A. Machine Learning-Based Prediction of Heating Values in Municipal Solid Waste. *Sci. Rep.* **2025**, *15*, 14589. [CrossRef]
71. Kujawska, J.; Kulisz, M.; Oleszczuk, P.; Cel, W. Improved Prediction of the Higher Heating Value of Biomass Using an Artificial Neural Network Model Based on the Selection of Input Parameters. *Energies* **2023**, *16*, 4162. [CrossRef]

72. Jayapal, A.; Morales, F.O.; Ishtiaq, M.; Kim, S.Y.; Reddy, N.G.S. Modeling the Higher Heating Value of Spanish Biomass via Neural Networks and Analytical Equations. *Energies* **2025**, *18*, 4067. [[CrossRef](#)]
73. Liu, J.; Iwakin, O.; Romero, C.E.; Cheng, L.; Moazeni, F.; Yao, Z.; De Saro, R.; Craparo, J. Rapid Characterization of MSW and RDF Feedstocks for Waste-to-Energy Process Using LIBS and ML Techniques. *Waste Manag.* **2025**, *206*, 115079. [[CrossRef](#)] [[PubMed](#)]
74. Taki, M.; Rohani, A. Machine Learning Models for Prediction the Higher Heating Value (HHV) of Municipal Solid Waste (MSW) for Waste-to-Energy Evaluation. *Case Stud. Therm. Eng.* **2022**, *31*, 101823. [[CrossRef](#)]
75. Güleç, F.; Pekaslan, D.; Williams, O.; Lester, E. Predictability of Higher Heating Value of Biomass Feedstocks via Proximate and Ultimate Analyses—A Comprehensive Study of Artificial Neural Network Applications. *Fuel* **2022**, *320*, 123944. [[CrossRef](#)]
76. Wang, X.; Zhao, S. The PSO/GA/ANN Modeling and Prediction for the Higher Heating Values of Solid Fuels: The Machine Learning Approach. *Therm. Sci.* **2025**, *29*, 2881–2897. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.