FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 4, April, 2025, pp 285 – 294
DOI: https://doi.org/10.33003/fjs-2025-0904-3493

PERFORMANCE COMPARISON OF DIFFERENT ACTIVATION FUNCTIONS IN NEURAL NETWORKS FOR BIOMASS ENERGY CONTENT PREDICTION

***Samuel Anjikwi Msheliza and Usman Alhaji Dodo**

Department of Electrical and Computer Engineering, Faculty of Engineering, Baze University, Abuja, Nigeria

*Corresponding authors' email: samule4899@bazeuniversity.edu.ng

ABSTRACT

The application of artificial neural networks to solve complex linear and nonlinear problems such as predicting biomass higher heating value (HHV) requires meticulous choice of the number of layers and neurons per layer, the training algorithms, and the activation functions among other hyper-parameters. Although some studies have examined these hyper-parameters, the effects of different activation functions on biomass HHV prediction have not attracted credible research attention. This study, therefore, employs three distinct activation functions (logsig, tansig, and purelin) in artificial neural networks for biomass HHV prediction based on proximate analysis. A 3-10-1 network architecture was used and the variation of the hidden layer and output layer activation functions yielded nine models (M1-M9) whose performances were assessed and compared using statistical indices. The results showed that the best performance was observed by model M2 which utlized the logsig function in the hidden layer and the tansig function in the output layer. This model had the highest determination coefficient of 0.8814 and the lowest mean square error and mean absolute error of 0.0017 and 0.0281 respectively. Understanding how these hyper-parameters influence biomass HHV prediction would guide the energy community to identify an optimal pathway to bioenergy production.

Keywords: Artificial neural networks, Biomass, Higher heating value, Proximate analysis, Transfer function

INTRODUCTION

The increase in population is a global issue that gives rise to various phenomena, each having adverse effects on the world at large, one of which is increased fuel consumption on a substantial scale. The increasing world population has resulted in substantial global emissions, which can equally be attributed to rapidly advancing industry and excessive reliance on fossil fuels to generate electricity and heat. Thus, the need for alternative sources of fuel is pressing, now more than ever (Dashti et al., 2019; Dodo et al., 2024).

Renewable energy sources, being sustainable and safe for the environment, are employed to slow down the rate at which fossil fuels are consumed. They include wind, solar, biomass, hydro, and tidal. Because of their availability and anticipated advantages for the local economy and environment, biomass, which can be defined as any combustible material, useable as fuels for the production of energy (heat and electricity) has continued to gain novel interest (Kujawska et al., 2023; Suryadevara et al., 2021). Thus, a variety of thermal, chemical, and biochemical processes can transform biomass into various types of biofuels, or it can be used directly in burners to produce heat and electricity (Dashti et al., 2019).

Certain biomass characteristics are crucial to the design and functioning of energy generation or energy recovery systems derived from biomass. The most notable characteristic is the energy content of biomass, also referred to as the heating value. The heating value is described as the total heat liberated when one unit mass of fuel is burned entirely, including the latent heat stored in the vapourised water as the product of combustion. Therefore, fuel with greater heating value will produce relatively higher energy output (Dodo, Ashigwuike, & Abba, 2022; Jakšić et al., 2023). Heating value is reported in the literature as either lower heating value (LHV) or higher heating value (HHV). Basically, the HHV comprises heat of vaporization while the LHV excludes it when a unit mass of fuel is entirely combusted. The most widely used experimental method to determine the heating value involves the application of an adiabatic oxygen bomb calorimeter which measures the enthalpy change between the reactants and products. Meanwhile, despite its reliability, this approach has the drawbacks of being time-consuming and expensive (Adeleke et al., 2024; Yang et al., 2023).

To address the drawbacks of the experimental procedures, alternative approaches that employ empirical models or artificial intelligence to predict the heating values of biomass have been developed by researchers. These techniques leverage either the proximate analysis or the ultimate analysis of the biomass substrates. However, biomass characteristics are nonlinear, making the empirical models incapable of capturing the features of the substrates entirely (Afolabi et al., 2022; Dodo et al., 2024). Hence, the use of artificial intelligence techniques such as artificial neural networks, adaptive neuro-fuzzy inference systems, decision trees, etc. Thus, they possess the ability to effectively describe and capture complicated and complex features of any system (Veza et al., 2022). Furthermore, the proximate analysis-based models are more widely utilized because of their lesser cost implications and ease of determination compared to the ultimate analysis-based models, (Veza et al., 2022; Xing et al., 2019). The proximate analysis utilizes known percentages of the fixed carbon (FC), volatile materials (VM), and ash. Naturally, these percentage residues from the combustion of the biomass must sum up to a hundred (FC+VM+ash=100) (Jakšić et al., 2023).

Artificial neural network (ANN) is an artificial intelligence technique for modeling linear and nonlinear variables and it has witnessed increased interest in applications regarding biomass research over time. Studies applying ANN for biomass HHV prediction employ either the proximate analysis approach, the ultimate analysis approach, or a combination of both. For example, Dodo et al. (2023) used an Artificial Neural Network (ANN) to develop proximate analysis-based biomass HHV prediction models, Brandić et al. (2022) employed an ultimate analysis approach to estimate the HHV of miscanthus using ANN, and Güleç et al. (2022) used ANN models trained by the combination of ultimate and proximate analyses. The data samples or species are another factor that influences the outcome of HHV prediction. For

instance, Yang et al. (2023) used ANN to predict the HHV of sewage sludge whereas Güleç et al. (2022), employed a dataset covering commercial fuels, industrial wastes, forest wastes (including branches and leaves) as well as energy crops and cereals. These studies have the limitations of being unsuitable for a broad range of biomass substrates such as agro residues and animal wastes.

The architecture of the ANN has a great effect on the prediction accuracy even more than the number of data samples (Dodo, Ashigwuike, & Abba, 2022). Hence, researchers applying ANN for the prediction of biomass HHV have used various network architectures. This includes the selection of an appropriate number of layers and hidden neurons, the training algorithms, and activation functions. In the work of Dodo et al. (2022), the relation $n + 1$ was applied to determine the number of hidden layers where n stands for the number of input parameters, and (Matveeva & Bychkov 2022) employed a trial-and-error method to select the optimal number of hidden layers between 1 and 6. The results of the study showed that overfitting occurs earlier as the number of hidden layers increases from one to six. In the same vein, Dodo et al. (2022) varied the number of hidden layer neurons from 1 and 20 using trial and error. The best performance was observed for the model with 15 hidden neurons using the trainbr algorithm which had a Nash-Sutcliffe's efficiency (NSE) value of 0.9044.

There are a handful of studies that explored different training algorithms to predict the biomass HHV. Güleç et al. (2022) performed a comparative analysis of thirteen different algorithms in ANN and found Levenberg-Marquardt (LM) and Bayesian regulation (BR) algorithms as the most suitable for HHV prediction. Veza et al. (2022) and Jakšić et al. (2023) developed ANN models trained using 11 and 12 different algorithms, respectively. They both ranked LM as the most accurate training algorithm in terms of its highest determination coefficient and the lowest mean square errors, root mean square error, and mean absolute percentage errors.

One notable aspect of ANN modelling that has not attracted due diligence is the selection of optimal activation functions for hidden and output layers (Balarabe et al., 2019; Pokhrel, 2024). Activation functions also known as the transfer functions define the properties of artificial neurons and can be any mathematical function. In a simple term, it is a mathematical depiction of the relationship between the input and output expressed in terms of spatial or temporal frequency (López et al., 2022). The functions which can be a step function, linear function, and nonlinear function (sigmoid) are chosen based on the problems that ANN needs to solve. A recent study by Laabid et al. (2023) examined the effects of three different activation functions on the response time of an ANN model. A 2-15-1 ANN architecture was employed while the hidden layer and output layer activation functions, respectively were selected from tansig, logsig, and purelin leading to nine different models. Thus, in every three models, the activation function of the hidden layer and that of the output layer were modified. The results showed the best performance belonging to the model which had the tansig function in both the hidden layer and the output layer. This study, however, used a dataset related to the elasticity of two distinct materials and not biomass HHV. In a nutshell, it is apparent from the review of related studies that the comparative study of different activation functions in artificial neural networks for biomass HHV prediction is limited. The activation function is a key component of neural networks that enables the network to learn and identify intricate patterns in input. Selecting the right activation function is crucial because it can influence the network's capacity to gather data and stop input data loss during forward propagation and gradient vanishing during backward propagation. Thus, it is vital to evaluate the prevalence of using different activation functions for biomass HHV prediction.

Therefore, this study aims to compare the performance of three different activation functions in feedforward backpropagation neural networks (FBNN) for biomass energy content prediction. The FBNN architecture implemented consists of three inputs from the proximate analysis, one output (HHV), and one hidden layer comprising ten hidden neurons. The three activation functions, namely purelin, logsig, and tansig were varied between the hidden layer and the output layer to identify the function capable of providing optimal performance. It is hoped that this study will serve as a paradigm for determining the energy content (HHV) of biomass economically and accurately.

MATERIALS AND METHODS

This study performed the comparative performance analysis of three activation functions, namely, tangent sigmoid, purelin, and log-sigmoid in FBNN for biomass HHV prediction based on proximate analysis. The HHV in MJ/kg and the proximate analysis variables comprising volatile matter (VM), fixed carbon (FC), and ash in wt.% were obtained from the published literature works (Estiati et al., 2016; Gunamantha, 2016; Nhuchhen & Salam, 2012; Phichai et al., 2013; Qian et al., 2018; Uzun et al., 2017).

Data pre-processing

The experimental datasets made of 474 instances of the HHV, FC, VM, and ash were normalized following eq. (1) to a range of 0–1 to enhance pattern recognition and the ANN's capacity to deliver superior results.

$$X_n = \frac{X_i - X_{min}}{Xmin_{max}} \quad (1)$$

where X_n is the normalised experimental parameter, X_i is the original experimental parameter, X_{max} and X_{min} are the maximum and minimum values of the experimental datasets respectively.

Next, an FBNN architecture for prediction was carefully developed after which the three activation functions were varied between the hidden and output layers, giving rise to nine models. For training, testing, and validation, the ANN toolboxes and their MATLAB counterparts use a 60:20:20 data division. According to the standard calibration of training-testing data, this could be folded into 60:40. The most prevalent data calibration in the scientific literature is either 70:30 or 75:25 based on the assumption that prediction models operate better with a bigger pool of training data (Dodo et al., 2024). Therefore, in this study, the partitioning applied was 80:20 for training and testing data respectively. The performances of the models developed were then assessed using a correlation-based metric and two error-based metrics. The steps followed for the implementation of the study are presented in a simple flowchart in Fig 1.

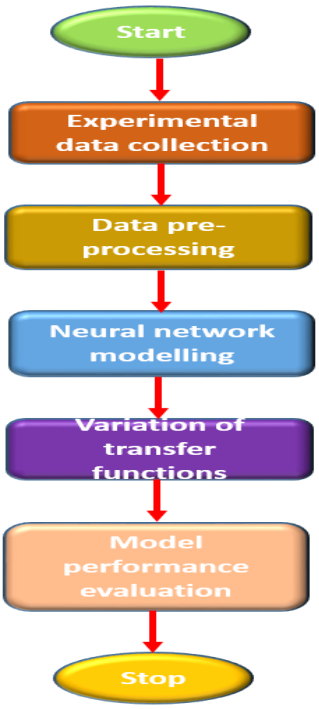


Figure 1: Flowchart of the methodology

FFNN implementation

An artificial neural network is a biologically inspired computer algorithm based on machine learning (Qamar & Baqar, 2023; Shehu & Belgore, 2023). In artificial neural networks, there are two types of connections between nodes known as the network topology. One is a one-way connection with no loopback. The other is a loop-back connection in which the output of the nodes can be the input to previous or same-level nodes. Based on the type of connections, the two types of network topology are feedforward networks and recurrent neural networks (Dastres et al., 2021), as shown in Fig. 2.

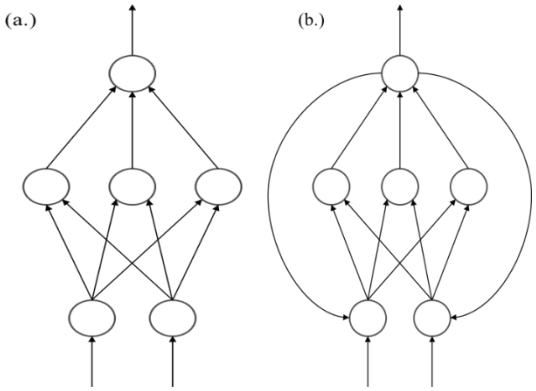


Figure 2: (a) Feedforward network topology (b) Recurrent network topology

In this study, the feed-forward backpropagation neural network (FFNN) was used due to its simplicity and effectiveness. The network design, including the number of hidden layers and neurons in each layer, the training algorithm, and the activation functions must be specified by the user when building the FFNN model. The neural network configuration used in this study is 3-10-1 as depicted in Fig. 3. It comprised three input neurons for FC, VM, and ash, ten neurons for a hidden layer (HN1-HN10), and a neuron for the output variable (HHV). The training algorithm employed was Levenmberg-Marquadrth (trainlm) due to its exceptional predictive performance (Dodo et al., 2023). The activation functions considered in this study include logarithmic sigmoid (logsig), hyperbolic tangent sigmoid (tansig), and linear (purelin).

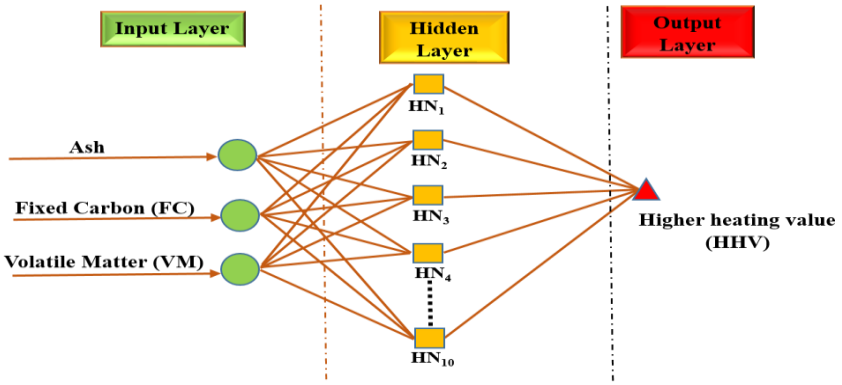


Figure 3: Schematic of a 3-10-1 FFNN architecture

Activation functions

An activation function also called transfer function is a mathematical depiction of the relationship between the input and output expressed in terms of spatial or temporal frequency (López et al., 2022). Typically, three traditional differentiable and monotonic activation functions are used for the evolution of the FFNN architecture in conjunction with the LM learning algorithm. These suggested, well-known, and efficient activation functions are logarithmic-sigmoid (logsig), hyperbolic tangent-sigmoid (tansig), and linear (purelin) (Nandi et al., 2020).

Purelin

The Purelin is a linear function that is usually applied to tasks involving regression and function approximation and is commonly used in the output layer of a neural network (Reyes-Téllez et al., 2020). It is expressed mathematically in eq. (2) and graphically in Fig. 4.

$$F(x) = Purelin(x) = x \quad (2)$$

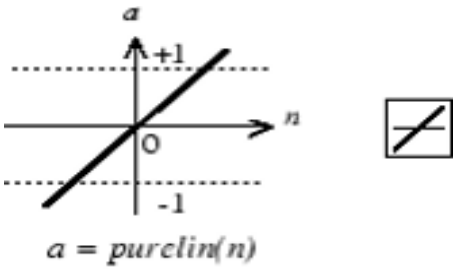


Figure 4: Purelin activation function

Log-sigmoid

Log-sigmoid (logsig) can also be referred to as a unipolar sigmoid activation function. This function in mathematics produces a sigmoidal curve, which is a typical curve for its S-shape (Lederer, 2021; Rasamoelina et al., 2020). This activation function takes the input which may have values between plus and minus infinity and compresses the output into the range 0 to 1. It is frequently employed in multilayer networks trained using the backpropagation process, partly because of the function's differential nature. Mathematically, it is expressed in eq. (3) and its graph is shown in Fig. 5.

$$F(x) = log\,sig(x) = (1 - e^{-x})^{-1} \quad (3)$$

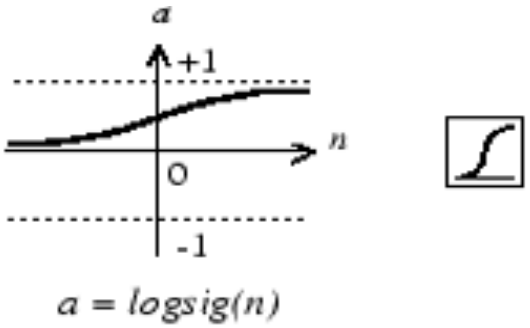


Figure 5: Log-sigmoid activation function

Tangent sigmoid

The Hyperbolic tangent sigmoid (tansig) activation function is comparable to a bipolar sigmoid in the context of neural networks, with an output that falls between -1 and +1 (Lederer, 2021; Rasamoelina et al., 2020). Mathematically, this function is expressed as the ratio between the hyperbolic sine and cosine functions, which is a hyperbolic tangent as seen in eq. (4) and its graph is depicted in Fig. 6.

$$F(x) = tanh(x) = \frac{2}{1+e^{-2x}} - 1 \quad (4)$$

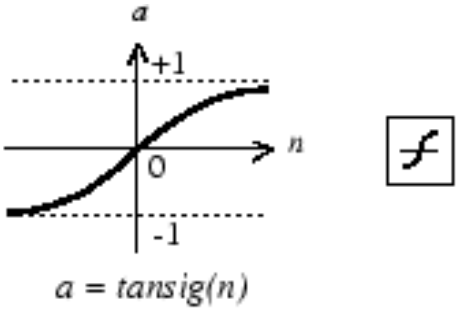


Figure 6: Tan-Sigmoid activation function

For the hidden layer and the output layer, the activation function was varied to find the combination that yielded the best results. The different combinations of the three activation functions across the hidden layer and the output layer yielded nine models labeled M1-M9 as shown in Table 1.

Table 1: Models based on the variation of activation function

FFNN Model	Hidden layer activation function	Output layer activation function
M1	Logsig	Logsig
M2	Logsig	Tansig
M3	Logsig	Purelin
M4	Tansig	Logsig
M5	Tansig	Tansig
M6	Tansig	Purelin
M7	Purelin	Logsig
M8	Purelin	Tansig
M9	Purelin	Purelin

Performance evaluation

An assessment of the performance of the M1-M9 for the training and testing phases was carried out by employing three evaluation metrics, namely, determination coefficient (R^2), mean square error (MSE), and mean absolute error (MAE). These metrics are expressed in eqs. (5)-(7).

$$R^2 = 1 - \frac{\left(HHV_{e(i)} - HHV_{p(i)}\right)^2}{\left(HHV_{e(i)} - \overline{HHV}_{e(i)}\right)^2} \quad (5)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(HHV_{e(i)} - HHV_{p(i)}\right)^2 \quad (6)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|HHV_{e(i)} - HHV_{p(i)}\right| \quad (7)$$

$\overline{HHV}_{e(i)}$, $HHV_{e(i)}$, $HHV_{p(i)}$,and$n$,respectively denote the mean of experimentalHHV, experimentalHHV, predictedHHV, and number of data instances.

RESULTS AND DISCUSSION

In this study, biomass HHV prediction models based on moisture-free proximate analysis (ash, VM, and FC) were developed using three activation functions varied between the hidden and output layers of an FFNN leading to nine different models (M1-M9). The aim was to examine which activation functions perform optimally for this application. The results in terms of the three evaluation metrics (R^2, MSE, and MAE) after varying the activation functions according to Table 1 were obtained for the training and testing phases and presented in Table 2. The visualizations and analyses carried out in this section cover both the training and testing phases.

Table 2: Models performance evaluation

FFNN Model	Training phase			Testing phase		
	R^2	MSE	MAE	R^2	MSE	MAE
M1	0.484434	0.007527	0.050395	0.316517	0.017584	0.081262
M2	0.881413	0.001731	0.028107	0.816413	0.004723	0.046794
M3	0.875093	0.001823	0.029103	0.841223	0.004085	0.042616
M4	0.472093	0.007707	0.054261	0.267929	0.018835	0.087050
M5	0.822980	0.002584	0.034771	0.825501	0.004489	0.050097
M6	0.843125	0.002290	0.031426	0.866297	0.003440	0.042605
M7	0.370370	0.009192	0.066920	0.284628	0.018405	0.088722
M8	0.814973	0.002701	0.037051	0.831285	0.004341	0.048668
M9	0.808911	0.002790	0.037646	0.843263	0.004032	0.047609

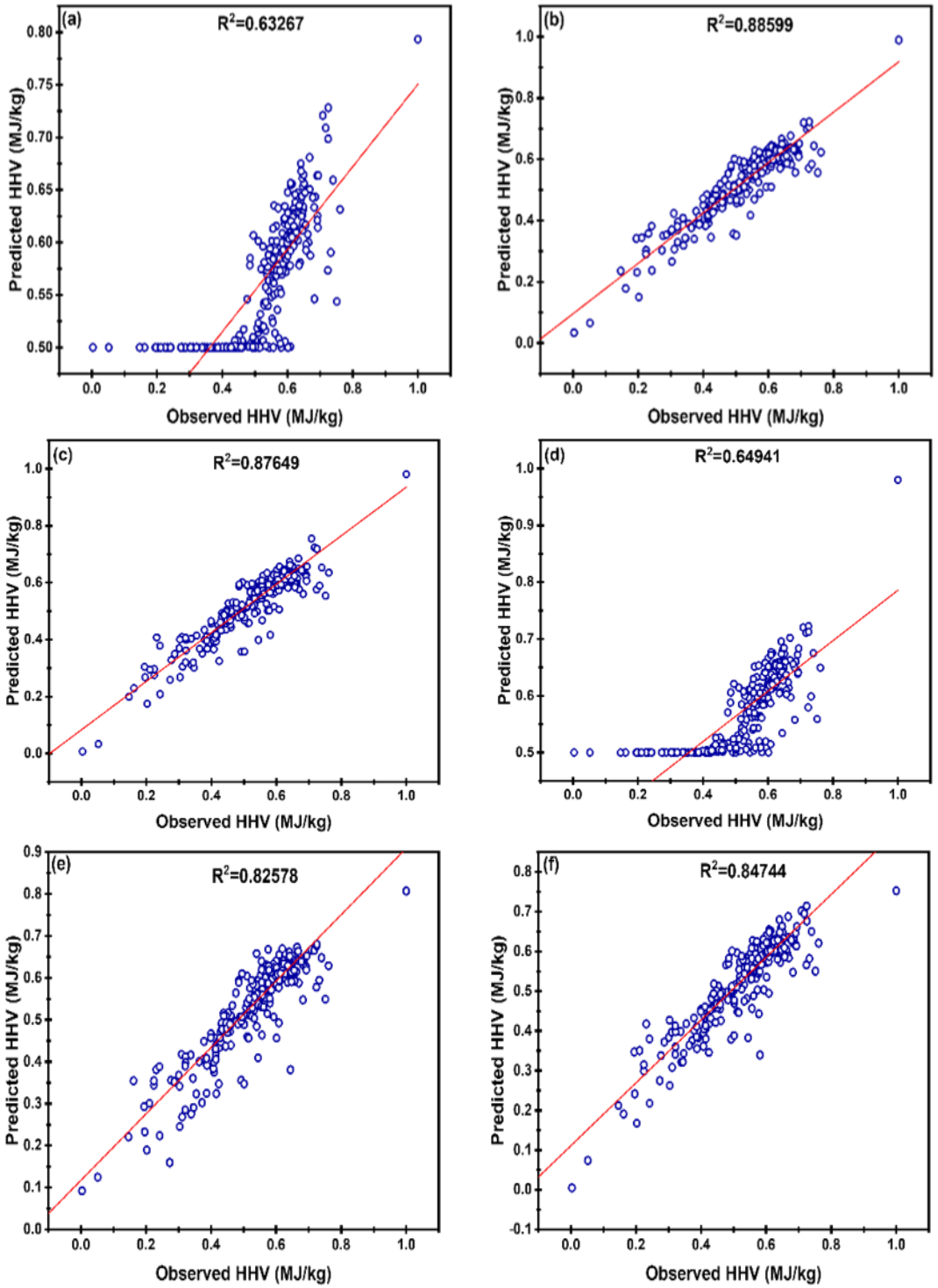
Prediction accuracy of the models

The models M1 to M9 were assessed based on three evaluation metrics: R^2, MSE, and MAE. Values obtained for the training and testing phases are presented in Table 2 and from an analysis of the table, the best-performing model, based on the results for the training phase can be identified as M2. This model was able to produce predicted HHVs which have a high correlation with experimental HHVs, hence the high coefficient of determination of 0.8814. The closer the value of the error metrics is to 0 the better the performance of the model, which means the low MSE and MAE values; 0.001731 and 0.0281 respectively, attained by model M2 are also indicators of good performance.

Model M7 however, had the weakest performance in the training phase. Generally, an R^2 value above or equal to 0.7 is considered highly satisfactory, while any value below that range is said to be unsatisfactory (Dodo et al., 2023). Model M7 had an R^2 value of 0.3704, which is marginally below the acceptable value, indicating a substantial underperformance and an incompatibility of the activation function combination used therein. The high MSE and MAE values of 0.009192 and 0.06692 respectively compared to M2 also point to the inefficiency of the model.

In the testing phase, the best model was M6 which achieved the highest R^2 value of 0.8663 and the lowest MSE and MAE values of 0.00344 and 0.04261 respectively. Meanwhile, model M4 had the poorest performance. The model had an R^2 value of 0.267929, the highest MSE value of 0.018835, and the second-highest MAE value of 0.08705.

Fig. 7 depicts scatter plots of experimental against predicted HHV for all nine models for a combination of the training and testing phase data. The models with good performances, hence, high R^2 values have the data points clustered close to the line of best fit while models with poor performance have data points scattered away from the line of best fit.



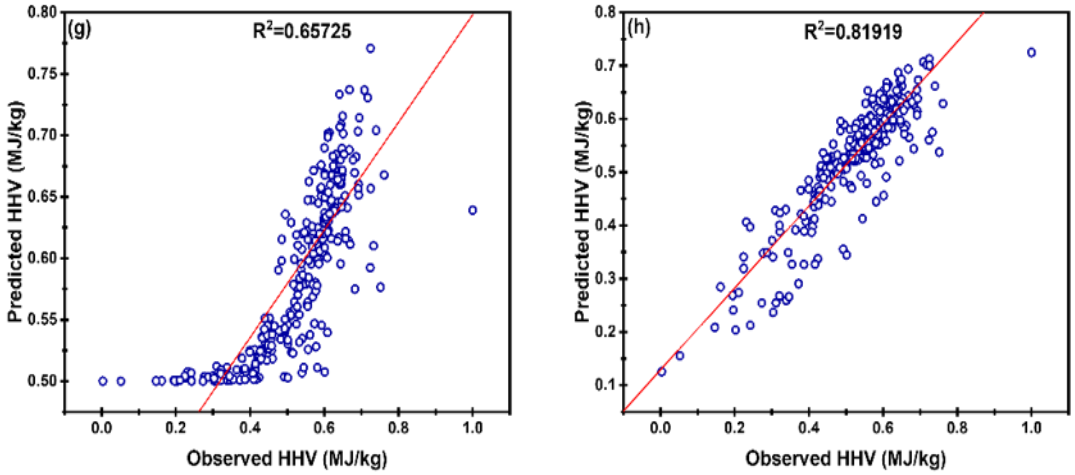


Figure 7: Scatter plots for various models (a) M1; (b) M2; (c) M3; (d) M4; (e) M5; (f) M6; (g) M7; (h) M8; (i) M9

Models performance comparison

Out of the nine models (M1-M9), six were able to accurately predict the HHV while three were inaccurate in prediction. The models M1, M4, and M7 were unable to deliver satisfactory results having training phase R^2 values of 0.4844, 0.472, and 0.3704 respectively with similarly low testing phase R^2 values of 0.3165, 0.2679, and 0.2846. These models have in common, the fact that they all utilized the logsig function in the output layer, which implies the function is not a good choice as the output layer activation function.

However, the opposite effect is noted when the logsig function was utilized in the hidden layer as evident from models M2 and M3. These models had the best and second-best performances respectively in terms of all three evaluation metrics used when considering the training phase. Similarly, when considering the testing data phase, the two models had good R^2, MSE, and MAE values in comparison with the other seven models that had the sixth and third-best performances respectively. Therefore, it is evident that the logsig function has exceptional performance in the hidden layer but produces unsatisfactory results when utilized in the output layer.

Models M6 and M5 were able to produce satisfactory results as well as had the third and fourth best training phase performances respectively in terms of the three evaluation metrics and also maintained similar comparative performance against other models in the testing phase. These two models utilized the tansig function in the hidden layer and their performance in both the training and testing phases leads to the inference that after the logsig function, the tansig performs best in the hidden layer. Finally, among the models which had accurate HHV predictions, models M8 and M9 had the fifth and sixth-best performances respectively when considering the training phase but had comparatively better performances when considering the testing phase where they had the second and fourth-best performances in terms of all three evaluation metrics. These models employed the purelin function in the input layer, thus, it can be inferred that the purelin function has the third-best performance among the three activation functions in the hidden layer.

The ranking of the models in the training phase is M2, M3, M6, M5, M8, M9, M1, M4, and M7 in descending order of prediction accuracy while the ranking of the models in the testing phase is M6, M9, M3, M8, M5, M2, M1, M7, and M4 in descending order of prediction accuracy. It is worth noting that logsig was the only activation function with unacceptable performance when utilized in the output layer and no other model recorded unsatisfactory performance except M1, M4 and M7 which had this common feature. This can be attributed to the fact that, unlike linear functions whose output has infinite range$(-\infty, \infty)$, the input values usually respond only slightly to changes in the input values near either end of the function, and sigmoidal functions compress the output values to a smaller range (Szandała 2021). This suggests that a slight gradient will exist at this location. As a result, a vanishing gradient issue arises, which is felt in the direction of the curve's near-horizontal activation functions on each side. Depending on how it is utilised until the gradient approaches the floating-point value restrictions, the network may either learn extremely slowly in this scenario or refuse to learn at all. This can be observed in scatter plots for models M1, M4, and M7 in Fig. 7.

The relative performance of each model in terms of R^2, MSE, and MAE are shown in Figs. 8-10. In the radar plot for R^2, the farther a point is from the centre, the higher the value, and the better the performance of that model. Hence it can be seen from Fig. 8 that models M2, M3, M5, M6, M8, and M9 had good prediction accuracy based on their high R^2 values for both training and testing phases. When considering error metrics like the MSE and MAE, however, the higher the value, the more inaccurate the model is. Therefore, from an examination of Figs. 9 and 10, it is evident that for both the training and testing phases, models M1, M4, and M7 do not deliver accurate results. Furthermore, the relationship between each of the predicted HHVs of each model and the experimental HHVs for the testing and training phases can be observed in Fig. 11. Variables having close symmetrical resemblance in a box plot indicate that the correlation between the values of those variables is high. In Fig. 11, models M2, M3, M5, M6, M8, and M9 have relatively similar symmetry to the HHV, indicating a good correlation between the predicted HHVs of these models and the experimental HHVs.

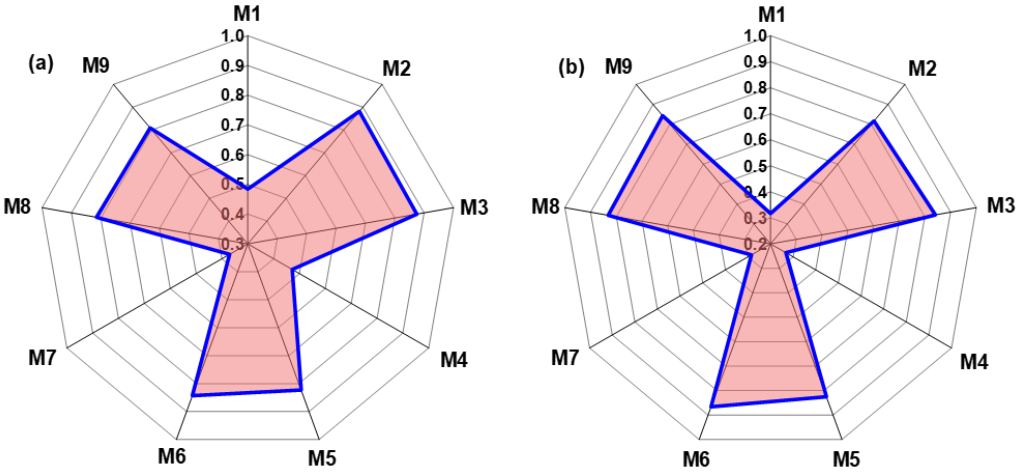


Figure 8: R^2 for M1 to M9 (a) Training phase (b) Testing phase

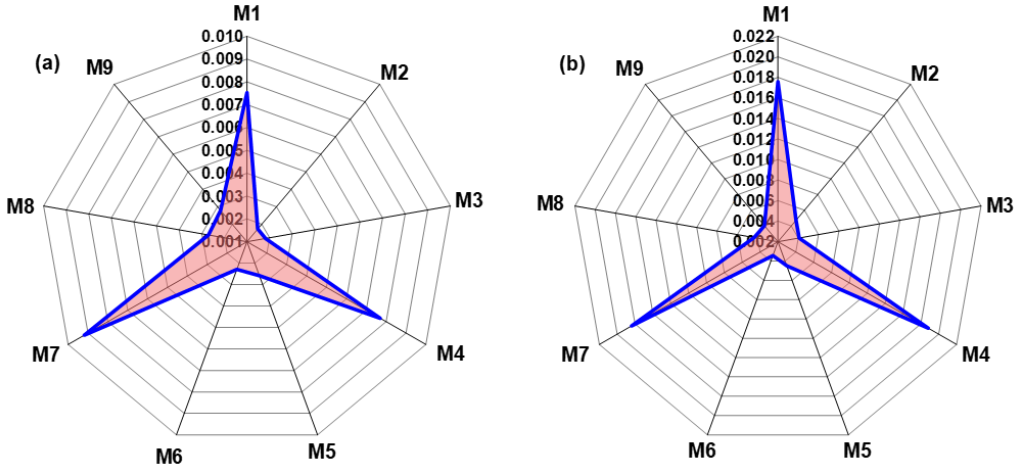


Figure 9: MSE for M1 to M9 (a) Training phase (b) Testing phase

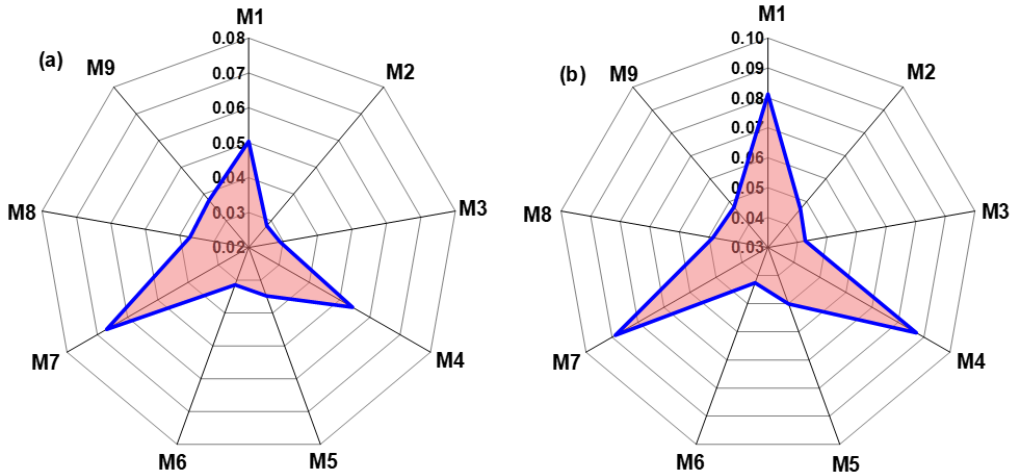


Figure 10: MAE for M1 to M9 (a) Training phase; (b) Testing phase

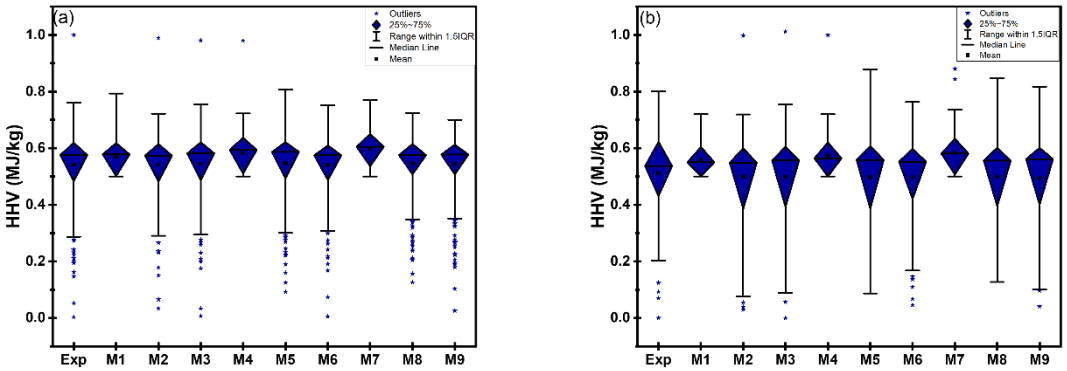


Figure 11: Predicted HHV for M1-M9 vs. Experimental HHV (a) Training phase (b) Testing phase.

CONCLUSION

In this study, a performance comparison of three different activation functions (logsig, tansig, and purelin) in FFNN to predict the HHV of a wide range of biomass substrates as a function of the proximate analysis was conducted. This was achieved by modifying the activation functions of the output and hidden layers of a 3-10-1 FFNN architecture, leading to a total of nine investigated models (M1-M9). The optimal performance is credited to model M2 which used the logsig function in the hidden layer and the tansig function in the output layer. This model had the highest R^2 value of 0.8814 and the lowest MSE and MAE values of 0.0017 and 0.0281 respectively. Meanwhile, the logsig activation function has the weakest performance when set as the output layer activation function but has the best performance when used as the hidden layer activation function based on all three evaluation metrics R^2, MSE, and MAE. This implies that, to implement FFNN for biomass HHV prediction using proximate analysis, the tansig should be prioritised as the output layer activation function followed by the purelin while for the hidden layer activation function, preference should be given to the logsig, then tansig and purelin in that order. Future studies may focus on employing these activation functions to predict biomass HHV as a function of the ultimate analysis.

REFERENCES

Adeleke, A. A., Adedigba, A., Adeshina, S. A., Ikubanni, P. P., Lawal, M. S., Olosho, A. I., Yakubu, H. S., Ogedengbe, T. S., Nzerem, P., & Okolie, J. A. (2024). Comparative studies of machine learning models for predicting higher heating values of biomass. *Digital Chemical Engineering*, *12*(100159), 10. https://doi.org/10.1016/j.dche.2024.100159

Afolabi, I. C., Epelle, E. I., Gunes, B., Güleç, F., & Okolie, J. A. (2022). Data-Driven Machine Learning Approach for Predicting the Higher Heating Value of Different Biomass Classes. *Clean Technologies*, *4*(4), 1227–1241. https://doi.org/10.3390/cleantechnol4040075

Balarabe, M. A., & Isah, M. N. (2019). Feed-Forward and Cascade Back Propagation Artificial Neural Network Models for. *FUDMA Journal of Sciences*, *3*(1), 428–433.

Brandić, I., Pezo, L., Bilandžija, N., Peter, A., & Šuri, J. (2022). *Artificial Neural Network as a Tool for Estimation of the Higher Heating Value of Miscanthus Based on Ultimate Analysis*. *10*(20), 1–12.

Dashti, A., Noushabadi, A. S., Raji, M., Razmi, A., Ceylan, S., & Mohammadi, A. H. (2019). Estimation of biomass higher heating value (HHV) based on the proximate analysis: Smart modeling and correlation. *Fuel*, *257*(115931), 1–11. https://doi.org/10.1016/j.fuel.2019.115931

Dastres, R., Soori, M., Neural, A., Systems, N., & Journal, I. (2021). Artificial Neural Network Systems To cite this version : HAL IAd : hal-03349542. *International Journal of Imaging and Robotics (IJIR)*, *21*(2), 13–25.

Dodo, U. A., Ashigwuike, E. C., & Abba, S. I. (2022). Machine learning models for biomass energy content prediction: A correlation-based optimal feature selection approach. *Bioresource Technology Reports*, *19*(101167), 11. https://doi.org/10.1016/j.biteb.2022.101167

Dodo, U. A., Ashigwuike, E. C., Emechebe, J. N., & Abba, I. S. (2022). Prediction of energy content of biomass based on hybrid machine learning ensemble algorithm. *Energy Nexus*, *8*, 15. https://doi.org/10.1016/j.nexus.2022.100157

Dodo, U. A., Dodo, M. A., Belgore, A. T., Husein, M. A., Ashigwuike, E. C., Mohammed, A. S., & Abba, S. I. (2024). Comparative study of different training algorithms in backpropagation neural networks for generalized biomass higher heating value prediction. *Green Energy and Resources*, *2*(1), 13. https://doi.org/10.1016/j.gerr.2024.100060

Dodo, U. A., Dodo, M. A., Shehu, A. F., & Badamasi, Y. A. (2023). Performance Analysis of Intelligent Computational Algorithms for Biomass Higher Heating Value Prediction. *Nigerian Journal of Technological Development*, *20*(4), 44–52. https://doi.org/10.4314/njtd.v18i4.1856

Estiati, I., Freire, F. B., Freire, J. T., Aguado, R., & Olazar, M. (2016). Fitting performance of artificial neural networks and empirical correlations to estimate higher heating values of biomass. *Fuel*, *180*, 377–383. https://doi.org/10.1016/j.fuel.2016.04.051

Güleç, F., Pekaslan, D., Williams, O., & Lester, E. (2022). Predictability of higher heating value of biomass feedstocks via proximate and ultimate analyses – A comprehensive study of artificial neural network applications. *Fuel*, *320*(123944), 1–16. https://doi.org/10.1016/j.fuel.2022.123944

Gunamantha, M. (2016). Prediction of Higher Heating Value Bioorganic Fraction of Municipal Solid Waste from Proximate Analysis Data. *International Journal of Engineering Research & Technology*, *5*(2), 442–447. http://www.ijert.org

Jakšić, O., Jakšić, Z., Guha, K., Silva, A. G., & Laskar, N. M. (2023). Comparing artificial neural network algorithms for prediction of higher heating value for different types of biomass. *Soft Computing*, *27*(9), 5933–5950. https://doi.org/10.1007/s00500-022-07641-4

Kujawska, J., Kulisz, M., Oleszczuk, P., & Cel, W. (2023). Improved Prediction of the Higher Heating Value of Biomass Using an Artificial Neural Network Model Based on the Selection of Input Parameters. *Energies*, *16*(10), 1–16. https://doi.org/10.3390/en16104162

Laabid, Z., Moumen, A., Mansouri, K., & Siadat, A. (2023). Numerical study of the speed's response of the various intelligent models using the tansig, logsig and purelin activation functions in different layers of artificial neural network. *IAES International Journal of Artificial Intelligence*, *12*(1), 155–161. https://doi.org/10.11591/ijai.v12.i1.pp155-161

Lederer, J. (2021). Activation Functions in Artificial Neural Networks: A Systematic Overview. In *Ruhr-University Bochum, Germany* (pp. 1–42). http://arxiv.org/abs/2101.09957

López, O. A. M., López, A., & Crossa, J. M. (2022). Fundamentals of Artificial Neural Networks and Deep Learning. In *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. https://doi.org/10.1007/978-3-030-89010-0_10

Matveeva, A., & Bychkov, A. (2022). How to Train an Artificial Neural Network to Predict Higher Heating Values of Biofuel. *Energies*, *15*(19), 1–13. https://doi.org/10.3390/en15197083

Nandi, A., Jana, N. D., & Das, S. (2020). Improving the Performance of Neural Networks with an Ensemble of Activation Functions. *Proceedings of the International Joint Conference on Neural Networks*, *December 2024*. https://doi.org/10.1109/IJCNN48605.2020.9207277

Nhuchhen, D. R., & Salam, P. A. (2012). Estimation of higher heating value of biomass from proximate analysis: A new approach. *Fuel*, *99*, 55–63. https://doi.org/10.1016/j.fuel.2012.04.015

Phichai, K., Pragrobpondee, P., Khumpart, T., & Hirunpraditkoon, S. (2013). Prediction Heating Values of Lignocellulosics from Biomass Characteristics. *International Journal of Chemical, Materials Science and Engineering*, *7*(7), 1–4.

Pokhrel, S. (2024). Intelligent Lighting Control Systems for Energy Savings in Hospital Buildings using Artificial Neural Networks. *FUDMA Journal of Sciences*, *15*(1), 37–48.

Qamar, R., & Baqar, A. Z. (2023). *Artificial Neural Networks : An Overview*. *2023*, 124–133.

Qian, X., Lee, S., Soto, A., & Chen, G. (2018). Regression Model to Predict the Higher Heating Value of Poultry Waste from Proximate Analysis. *Resources*, *7*(39), 1–14. https://doi.org/10.3390/resources7030039

Rasamoelina, A. D., Adjailia, F., & Sincak, P. (2020). A Review of Activation Function for Artificial Neural Network. *SAMI 2020 - IEEE 18th World Symposium on Applied Machine Intelligence and Informatics, Proceedings, July 2024*, 281–286. https://doi.org/10.1109/SAMI48414.2020.9108717

Reyes-Téllez, E. D., Parrales, A., Ramírez-Ramos, G. E., Hernández, J. A., Urquiza, G., Heredia, M. I., & Sierra, F. Z. (2020). Analysis of transfer functions and normalizations in an ann model that predicts the transport of energy in a parabolic trough solar collector. *Desalination and Water Treatment*, *200*, 23–41. https://doi.org/10.5004/dwt.2020.26063

Shehu, A. F., & Belgore, A. T. (2023). Machine Learning Approach to Wind Speed Prediction using Soft Computing Tools. *Journal of Science Technology and Education*, *11*(2), 349–355. www.atbuftejoste.net

Suryadevara, S., Kumar, A., & Yanamala, Y. (2021). *A Comprehensive Overview of Artificial Neural Networks : Evolution , Architectures , and Applications*. *12*(01), 51–76.

Szandała, T. (2021). Review and comparison of commonly used activation functions for deep neural networks. In *Studies in Computational Intelligence* (Vol. 903, pp. 203–224). https://doi.org/10.1007/978-981-15-5495-7_11

Uzun, H., Yıldız, Z., Goldfarb, J. L., & Ceylan, S. (2017). Improved prediction of higher heating value of biomass using an artificial neural network model based on proximate analysis. *Bioresource Technology*, *234*, 122–130. https://doi.org/10.1016/j.biortech.2017.03.015

Veza, I., Irianto, Panchal, H., Paristiawan, P. A., Idris, M., Fattah, I. M. R., Putra, N. R., & Silambarasan, R. (2022). Improved prediction accuracy of biomass heating value using proximate analysis with various ANN training algorithms. *Results in Engineering*, *16*(100688), 1–6. https://doi.org/10.1016/j.rineng.2022.100688

Xing, J., Luo, K., Wang, H., Gao, Z., & Fan, J. (2019). A comprehensive study on estimating higher heating value of biomass from proximate and ultimate analysis with machine learning approaches. *Energy*, *188*(116077), 1–35. https://doi.org/10.1016/j.energy.2019.116077

Yang, X., Li, H., Wang, Y., & Qu, L. (2023). Predicting Higher Heating Value of Sewage Sludges via Artificial Neural Network Based on Proximate and Ultimate Analyses. *Water (Switzerland)*, *15*(4), 16. https://doi.org/10.3390/w15040674