

Article

A New Method for Commercial-Scale Water Purification Selection Using Linguistic Neural Networks

Saleem Abdullah ^{1, }, Alaa O. Almagrabi ^{2,*} and Nawab Ali ¹

¹ Department of Mathematics, Abdul Wali Khan University Mardan, Mardan 23200, Pakistan; saleemabdullah@awkum.edu.pk (S.A.)

² Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

* Correspondence: aalmagrabi3@kau.edu.sa

Abstract: A neural network is a very useful tool in artificial intelligence (AI) that can also be referred to as an ANN. An artificial neural network (ANN) is a deep learning model that has a broad range of applications in real life. The combination and interrelationship of neurons and nodes with each other facilitate the transmission of information. An ANN has a feed-forward neural network. The neurons are arranged in layers, and each layer performs a particular calculation on the incoming data. Up until the output layer, which generates the network's ultimate output, is reached, each layer's output is transmitted as an input to the subsequent layer. A feed-forward neural network (FFNN) is a method for finding the output of expert information. In this research, we expand upon the concept of fuzzy neural network systems and introduce feed-forward double-hierarchy linguistic neural network systems (FFDHLNNS) using Yager–Dombi aggregation operators. We also discuss the desirable properties of Yager–Dombi aggregation operators. Moreover, we describe double-hierarchy linguistic term sets (DHLTSs) and discuss the score function of DHLTSs and the distance between any two double-hierarchy linguistic term elements (DHLTEs). Here, we discuss different approaches to choosing a novel water purification technique on a commercial scale, as well as some variables influencing these approaches. We apply a feed-forward double-hierarchy linguistic neural network (FFDHLNN) to select the best method for water purification. Moreover, we use the extended version of the Technique for Order Preference by Similarity to Ideal Solution (extended TOPSIS) method and the grey relational analysis (GRA) method for the verification of our suggested approach. Remarkably, both approaches yield almost the same results as those obtained using our proposed method. The proposed models were compared with other existing models of decision support systems, and the comparison demonstrated that the proposed models are feasible and valid decision support systems. The proposed technique is more reliable and accurate for the selection of large-scale water purification methods.

Keywords: double-hierarchy linguistic term set; Dombi t-norms; artificial neural network; decision-making

MSC: 94D05; 90C70



Citation: Abdullah, S.; Almagrabi, A.O.; Ali, N. A New Method for Commercial-Scale Water Purification Selection Using Linguistic Neural Networks. *Mathematics* **2023**, *11*, 2972. <https://doi.org/10.3390/math11132972>

Academic Editor: Francisco Rodrigues Lima Junior

Received: 5 April 2023

Revised: 16 May 2023

Accepted: 31 May 2023

Published: 3 July 2023



Copyright: © 2023 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

1.1. A Brief Review of the Development of Neural Networks and Their Types

Classical statistical methods have been widely used in various industries for decades, particularly in fields such as quality control, experimental design, and process optimization. However, in recent years, neural networks (NNs) [1] have emerged as powerful tools for solving complex problems in various fields, including finance, engineering [2], medicine [3], and computer science [4]. NNs have gained popularity due to their ability to handle large and complex data sets, learn patterns and relationships in the data, and make accurate predictions or classifications. In comparison to classical statistical methods [5], NNs have the

advantage of being able to model nonlinear relationships and capture complex interactions between variables. In the field of pattern recognition, NNs have been shown to outperform classical statistical methods, particularly when dealing with complex and high-dimensional data. In prediction and classification tasks, NNs have also been successful, achieving high accuracy rates in fields such as image [6,7] and speech recognition [8], natural language processing [9,10], and sentiment analysis [11,12]. Overall, while classical statistical methods remain useful in many applications, NNs have become a popular and powerful tool for solving complex problems in various fields. Regression [13] and NNs are often seen as competing model-building methods, as they can both be used for modeling and predicting outcomes based on input variables. However, while regression is a linear method that requires the assumption of linearity between the input and output variables, NNs are capable of modeling nonlinear relationships and do not require such assumptions. The structure and operation of the brain served as the first inspiration for neural networks, which are created to resemble the behavior of organic neurons. As a result, they share some performance characteristics with human neural biology. Layers of linked nodes or neurons that are arranged into input, hidden, and output layers are used in NNs to replicate the behavior of the human brain. The input layer receives data and sends them to the hidden layers for processing, while the output layer produces the final result. In order to increase the model's accuracy, the hidden layers are in charge of observing patterns and correlations in the data and modifying the weights of the connections between neurons. One of the key features of neural networks is their ability to learn from data and store that knowledge in the form of learned parameters. This allows them to make predictions [14,15] and classifications [16,17] based on previously seen examples and to generalize that knowledge to new, unseen examples. Neural networks are also capable of identifying patterns [18] in data, even in the presence of noise [19] or other sources of variability. This makes them useful for a variety of applications, including natural language processing, image and speech recognition, and predictive modeling [20,21]. In addition, neural networks are capable of taking past experiences into consideration and using that information to make inferences and judgments about new situations. This is known as "contextual learning," and it allows neural networks to adapt to changing conditions and make more accurate predictions over time. Overall, NNs are a powerful tool for processing and analyzing complex data, and their performance characteristics make them well-suited to a wide range of tasks in fields such as machine learning [22], artificial intelligence [23,24], and cognitive science [25]. There are different kinds of NNs, each with its own unique architecture and characteristics. Among the most common types, we can find feed-forward neural networks [26], recurrent neural networks (RNNs) [27], convolutional neural networks (CNNs) [28], auto-encoder neural networks [29], generative adversarial networks [30], etc. In this paper, we discuss feed-forward neural networks in detail.

1.2. A Brief Review of Feed-Forward Neural Networks and Their Uses

The connections between the nodes in feed-forward neural networks (FFNNs) do not cycle, and data only move from the input layer to the output layer in one direction. This makes FFNNs a relatively simple type of neural network compared to others such as RNNs and CNNs. Numerous applications have made use of FFNNs, such as natural language processing [10,11], image [7,8] and speech recognition [9], and financial forecasting [31]. They have been among the most successful learning algorithms and have been the basis for many other types of neural networks. However, FFNNs are a method for determining experts' output, similar to techniques such as the Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) method [32] and the grey relational analysis (GRA) method [33], among others. FFNNs are machine learning models that learn to make predictions or classifications based on input data. They do not rely on explicit expert knowledge or rules but rather on patterns and relationships in the data themselves. TOPSIS and GRA, on the other hand, are decision-making methods used in multi-criteria decision analysis (MCDA) [34]. These methods involve comparing alternatives based on multiple

criteria and weighing their importance to reach a decision. They do not involve machine learning or neural networks.

1.3. A Brief Review of Activation Function and Its Importance

Both the output layer and hidden layers of a FFNN can use activation functions [35]. Activation functions are essential in neural networks because they introduce nonlinearity, which allows the network to simulate complicated interactions between input and output variables. The rectified linear unit (ReLU) [36] and the family of sigmoid functions (including the logistic sigmoid function [37], hyperbolic tangent [38], and arctangent function [38]) are popular activation functions used in FFNNs. ReLU is widely used in the hidden layers because of its simplicity and effectiveness in training deep networks. It outputs zero for negative input and increases linearly for positive input. On the other hand, the logistic sigmoid function is commonly used in the output layer to transform the output of the network into a probability between 0 and 1. The logistic sigmoid function has a smooth and differentiable curve, which makes it easy to compute the gradient during back-propagation. The logistic sigmoid function's output is constantly between 0 and 1, making it appropriate for binary classification tasks. However, it is susceptible to the vanishing gradient problem, particularly in deep neural networks [39], which can result in delayed training and poor performance. Overall, choosing the appropriate activation function for a neural network depends on the specific problem at hand and the structure of the network.

The main objectives of this research are:

1. There are numerous aggregation operators, such as Einstein, Yager, and Dombi. In this research, we combine the Yager [40–42] and Dombi [43] aggregation operators to make a new aggregation operator called the Yager–Dombi aggregation operator and explain its desirable properties;
2. We expand the concept of a feed-forward neural network to incorporate a feed-forward double-hierarchy linguistic neural network using Yager–Dombi operators;
3. We develop a fuzzy neural network for the selection of water purification methods using a double-hierarchy linguistic neural network and use it for the selection of water purification methods;
4. We extend the Yager–Dombi operations to aggregate a double hierarchy for fuzzy information.

1.4. Motivation behind the Study

According to the study analysis above, there has been no extensive usage of double-hierarchy linguistic term sets (DHLTSs) in the field of FFNNs or of the Yager–Dombi aggregation operator.

The primary objectives of this study are as follows:

1. To extend the concept of fuzzy neural networks to incorporate double-hierarchy linguistic neural networks;
2. To combine existing aggregation operators to create a new aggregation operator;
3. To develop a fuzzy neural network for the selection of water purification methods;
4. To extend the Yager–Dombi operations to aggregate double-hierarchy fuzzy information.

1.5. Contribution of the Study

In this paper, we combine two t-norms (Yager and Dombi t-norms) and apply them to linguistic neural networks, develop a new model of linguistic neural networks, and solve the selection problem of the water purification procedure. The contribution of this paper can be summarized as follows:

1. We develop new t-norms and their operations by using the Yager and Dombi t-norms and discuss their relationships;
2. The developed t-norms are further expanded to aggregation operators to develop a new set of double-hierarchy linguistic terms;

3. The proposed aggregation is necessary for artificial neural networks. Therefore, we integrate the proposed aggregation operators into the hidden layers of a linguistic neural network;
4. We develop a new approach to linguistic neural networks and linguistic decision models using linguistic neural networks;
5. The proposed linguistic decision model, based on a linguistic neural network, is applied to water-purification procedure-selection problems;
6. The proposed models are verified and compared with other models in Sections 7 and 8 for validation.

The remaining sections of the paper are structured as follows: In Section 2, we discuss some fundamental definitions. In Section 3, we define the Yager–Dombi t-norm and t-conorm and discuss the DHLTYDWA, DHLTYDOWA, and DHLTYDHW operators and their desirable properties. We also discuss the score function and the distance between any two DHLTSs. In Section 4, we discuss activation functions and feed-forward neural networks. In Section 5, we discuss the output of the feed-forward neural network using the Yager–Dombi operator. Section 6 illustrates the selection of the best method for water purification. In Section 7, we use the extended TOPSIS [44] and GRA [45] methods for verification. In Section 8, we compare the extended TOPSIS and GRA methods with our proposed method and rank the output using different Yager–Dombi aggregation operators. In Section 9, the authors conclude their study and elaborate on its value and development direction.

2. Fundamental Concept

This section discusses fuzzy sets, intuitionistic fuzzy sets, a hierarchy linguistic term set, and a double-hierarchy linguistic term set.

If Y is a non-empty set, then an object having the form (L.ZADEH [46])

$$Z = \{ \langle y, U_Z(y) \rangle | y \in Y \} \tag{1}$$

is called a fuzzy set, and $U_Z(y) \in [0, 1]$ denotes the degree of membership of Y in Z . An object having the form

$$Z = \{ \langle y, U_Z(y), V_Z(y) \rangle | y \in Y \} \tag{2}$$

is called an intuitionistic fuzzy set, with $U_Z(y) \in [0, 1]$ denoting the degree of membership and $V_Z(y) \in [0, 1]$ representing the degree of non-membership of Y in Z . Additionally, the following condition is also satisfied:

$$U_Z(y) + V_Z(y) \in [0, 1],$$

Fuzzy sets are a mathematical framework for representing and manipulating uncertainty and imprecision in information. In contrast to traditional sets, which define membership in a binary, all-or-nothing way, fuzzy sets allow for partial membership, meaning that an element can belong to a set to a certain degree. The usefulness of fuzzy sets lies in their ability to model and reason about real-world problems that involve uncertainty, ambiguity, and vagueness. For example, in many domains such as decision-making, control systems, and artificial intelligence, it is often difficult or impossible to make precise distinctions between different categories or values. Fuzzy sets provide a flexible and intuitive way to represent and reason about such situations. They allow for a more nuanced and realistic representation of uncertainty and imprecision, which can lead to better decision-making and more robust system design. Several terms that define the variable in a natural language manner, such as “low”, “medium”, and “high”, compose a fuzzy linguistic term set. The degree to which a variable belongs to a given term is described by a fuzzy set that is assigned to each term. Let S be a non-empty set. Then, the linguistic term set is defined as:

$$S = \{ s_\alpha | \alpha \in [-\beta, \beta] \} \tag{3}$$

Suppose $S = \{s_\alpha | \alpha \in [-\beta, \beta]\}$ is a FHLT (first-hierarchy linguistic term set) and $O = \{o_p | p \in [-k, k]\}$ is a SHLTS (second-hierarchy linguistic term set), then the double-hierarchy linguistic term set (DHLTS) is defined as:

$$S_0 = \{S_\alpha \langle O_p \rangle | \alpha \in [-\beta, \beta], p \in [-k, k]\} \tag{4}$$

Gou et al. suggested two transformed functions between the DHLTs subscript and its numerical scale in order to deal with DHLTSs [47] more effectively. $S_0 = S_\alpha \langle O_p \rangle$ is an unbroken DHLTS. The transformed functions f and f^{-1} between the numerical value and the subscript $[\alpha, p]$ of the DHLT $S_\alpha \langle O_p \rangle$ are given below:

$$f : [-\beta, \beta] \times [-k, k] \rightarrow [0, 1],$$

$$= \frac{p + (\alpha + \beta)k}{2\beta k} = \Upsilon \tag{5}$$

$$f^{-1} : [0, 1] \rightarrow [-\beta, \beta] \times [-k, k],$$

$$f^{-1}(\Upsilon) = S_{[2\beta\Upsilon - \Upsilon]} \langle O_{k\{(2\beta\Upsilon - \Upsilon) - [2\beta\Upsilon - \Upsilon]\}} \rangle \tag{6}$$

where $[2\beta\Upsilon - \Upsilon]$ represents the integer part of the value $2\beta\Upsilon - \Upsilon$.

3. Yager–Dombi Operators for DHLTSs

This section gives the basic concepts of Yager–Dombi t-norm, Yager–Dombi t-conorm, Yager–Dombi operators, and Yager–Dombi aggregation operators for DHLTSs.

If G and H are any two numbers, then the Yager–Dombi t-norm and Yager–Dombi t-conorm are given as follows:

$$T(G, H) = \frac{1}{2 + \left\{ \left\{ \left\{ \frac{1 - \{1 - \min(1, (1-G)^t\}\}}{\min(1, (1-G)^t)} \right\}^k + \left\{ \frac{1 - \{1 - \min(1, (1-H)^t\}\}}{\min(1, (1-H)^t)} \right\}^k \right\}^{\frac{1}{t}} \right\}^{\frac{1}{k}}} \tag{7}$$

$$T''(G, H) = 1 - \frac{1}{1 + \left\{ \left\{ \left\{ \frac{\min(1, (G^t + H^t))}{1 - \min(1, (G^t + H^t))} \right\}^{\frac{k}{t}} + \left\{ \frac{\min(1, (G^t + H^t))}{1 - \min(1, (G^t + H^t))} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}}} \tag{8}$$

where $k, t \geq 1$ and $G, H \in [0, 1]$.

To define operation rules based on Yager–Dombi t-norms for this, let $S_1 = F(S_{\alpha_1} \langle O_{p_1} \rangle)$ and $S_2 = F(S_{\alpha_2} \langle O_{p_2} \rangle)$ be any two DHLTSs. Let $k, t \geq 1$, and $\beta > 0$ be any real numbers. Then, we establish the Yager–Dombi operators for DHLTSs, which are given below:

$$\text{I. } S_1 \oplus S_2 = 1 - \frac{1}{1 + \left\{ \left\{ \left\{ \frac{\min(1, F(S_{\alpha_1} \langle O_{p_1} \rangle))^t}{1 - \min(1, F(S_{\alpha_1} \langle O_{p_1} \rangle))^t} \right\}^{\frac{k}{t}} + \left\{ \frac{\min(1, F(S_{\alpha_2} \langle O_{p_2} \rangle))^k}{1 - \min(1, F(S_{\alpha_2} \langle O_{p_2} \rangle))^k} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}}} \tag{9}$$

$$\text{II. } S_1 \otimes S_2 = \frac{1}{2 + \left\{ \left\{ \left\{ \frac{1 - \{1 - \min(1, (1 - F(S_{\alpha_1} \langle O_{p_1} \rangle))^t)\}}{\min(1, (1 - F(S_{\alpha_1} \langle O_{p_1} \rangle))^t)} \right\}^k + \left\{ \frac{1 - \{1 - \min(1, (1 - F(S_{\alpha_2} \langle O_{p_2} \rangle))^t)\}}{\min(1, (1 - F(S_{\alpha_2} \langle O_{p_2} \rangle))^t)} \right\}^k \right\}^{\frac{1}{t}} \right\}^{\frac{1}{k}}} \tag{10}$$

$$\text{III. } \mathbb{B} \odot S_1 = 1 - \frac{1}{1 + \left\{ (\mathbb{B}) \left\{ \frac{\min(1, F(S_{\alpha 1} \langle O_{\mathbf{P}_1} \rangle))^t}{1 - \min(1, F(S_{\alpha 1} \langle O_{\mathbf{P}_1} \rangle))^t} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}}}$$

$$\text{IV. } (S_1)^{\mathbb{B}} = \frac{1}{2 + \left\{ (\mathbb{B}) \left\{ \left\{ \frac{1 - \{1 - \min(1, (1-G)^t\})\}^k}{\min(1, (1-G)^t)} \right\}^{\frac{1}{t}} \right\}^{\frac{1}{k}}}$$

If $S_1 = F(S_{\alpha 1} \langle O_{\mathbf{P}_1} \rangle)$ and $S_2 = F(S_{\alpha 2} \langle O_{\mathbf{P}_2} \rangle)$ are any two DLTEs, then the distance between $S_1 = F(S_{\alpha 1} \langle O_{\mathbf{P}_1} \rangle)$ and $S_2 = F(S_{\alpha 2} \langle O_{\mathbf{P}_2} \rangle)$ is defined as:

$$d(S_1, S_2) = \left| F(S_{\alpha 1} \langle O_{\mathbf{P}_1} \rangle) - F(S_{\alpha 2} \langle O_{\mathbf{P}_2} \rangle) \right| \tag{9}$$

where $d(S_1, S_2) \in [0, 1]$.

Suppose $S_i = \{F(S_{\alpha i} \langle O_{\mathbf{P}_i} \rangle) | i = 1, 2, 3 \dots n\}$ is a group of DHLTSs, then the double-hierarchy linguistic term Yager–Dombi weighted averaging (DHLTYDWA) operator is a mapping $\mathbb{Q}^n \rightarrow \mathbb{Q}$ such that:

$$DHLTYDWA(S_1, S_2, S_3, S_4, \dots, S_i) = \bigoplus_{i=1}^n W_i S_i$$

$W_i = (W_1, W_2, W_3, W_4 \dots W_i)^t$ is the weight vector of $S_i = \{F(S_{\alpha i} \langle O_{\mathbf{P}_i} \rangle) | i = 1, 2, 3 \dots n\}$ that fulfills the condition that $\sum_{i=1}^n W_i = 1$.

In the following theorem, we prove that the aggregated values obtained from different DHLTSs by using the proposed aggregation operators are again DHLTSs. This means that the aggregation operators are valid. Additionally, we provide validation for the proposed aggregation operators using Yager–Dombi t-norms.

Theorem 1. Suppose $S_i = \{F(S_{\alpha i} \langle O_{\mathbf{P}_i} \rangle) | i = 1, 2, 3 \dots n\}$ is a group of DHLTSs, then the aggregated value of DHLTSs using the DHLTYDWA operator is also a DHLTS, that is:

$$DHLTYDWA(S_1, S_2, S_3, S_4, \dots, S_i) = \bigoplus_{i=1}^n W_i S_i = 1 - \frac{1}{1 + \left\{ (\sum_{i=1}^n W_i) \left\{ \frac{\min(1, F(S_{\alpha i} \langle O_{\mathbf{P}_i} \rangle))^t}{1 - \min(1, F(S_{\alpha i} \langle O_{\mathbf{P}_i} \rangle))^t} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}}}$$
 \tag{10}

$W_i = (W_1, W_2, W_3, W_4 \dots W_i)^t$ is the weight vector of $S_i = \{F(S_{\alpha i} \langle O_{\mathbf{P}_i} \rangle) | i = 1, 2, 3 \dots n\}$ that fulfills the condition that $\sum_{i=1}^n W_i = 1$.

Proof. To prove this theorem, we employ the mathematical induction approach.

For $n = 1$, we have:

$$DHLTYDWA(S_1) = W_1 S_1 = 1 - \frac{1}{1 + \left\{ (W_1) \left\{ \frac{\min(1, F(S_{\alpha 1} \langle O_{\mathbf{P}_1} \rangle))^t}{1 - \min(1, F(S_{\alpha 1} \langle O_{\mathbf{P}_1} \rangle))^t} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}}}$$

As a result, the statement holds true for $n = 1$.

Assume that the above result holds true for $n = h$.

$$DHLTYDWA(S_1, S_2, S_3, \dots, S_h) = \bigoplus_{i=1}^h W_i S_i = 1 - \frac{1}{1 + \left\{ \left(\sum_{i=1}^h W_i \right) \left\{ \frac{\min(1, F(S_{\alpha_i} \langle O_{P_i} \rangle))^t}{1 - \min(1, F(S_{\alpha_i} \langle O_{P_i} \rangle))^t} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}}}$$

Next, we are going to show that the result is true for $n = h + 1$:

$$\begin{aligned} DHLTYDWA(S_1, S_2, S_3, S_4, \dots, S_{h+1}) &= \bigoplus_{i=1}^{h+1} W_i S_i \\ &= \left(1 - \frac{1}{1 + \left\{ \left(\sum_{i=1}^h W_i \right) \left\{ \frac{\min(1, F(S_{\alpha_i} \langle O_{P_i} \rangle))^t}{1 - \min(1, F(S_{\alpha_i} \langle O_{P_i} \rangle))^t} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}} \right) \oplus \left(1 - \frac{1}{1 + \left\{ W_{h+1} \left\{ \frac{\min(1, F(S_{\alpha_{h+1}} \langle O_{P_{h+1}} \rangle))^t}{1 - \min(1, F(S_{\alpha_{h+1}} \langle O_{P_{h+1}} \rangle))^t} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}} \right) \\ &= 1 - \frac{1}{1 + \left\{ \left(\sum_{i=1}^{h+1} W_i \right) \left\{ \frac{\min(1, F(S_{\alpha_i} \langle O_{P_i} \rangle))^t}{1 - \min(1, F(S_{\alpha_i} \langle O_{P_i} \rangle))^t} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}}} \end{aligned}$$

The result is valid for $n = h + 1$. Moreover, the DHLTYDWA operator can easily hold its idempotence, boundedness, and monotonicity properties.

Suppose $S_i = \{F(S_{\alpha_i} \langle O_{P_i} \rangle) | i = 1, 2, 3 \dots n\}$ is a group of DHLTSs, then the double-hierarchy linguistic term Yager–Dombi order weighted averaging (DHLTYDOWA) operator is a mapping $\mathbb{Q}^n \rightarrow \mathbb{Q}$ such that:

$$DHLTYDOWA(S_1, S_2, S_3 \dots S_i) = \bigoplus_{i=1}^n W_i S_{R(i)}$$

$W_i = (W_1, W_2, W_3, W_4 \dots W_i)^t$ is the weight vector of $S_i = \{F(S_{\alpha_i} \langle O_{P_i} \rangle) | i = 1, 2, 3 \dots n\}$ that fulfills the condition that $\sum_{i=1}^n W_i = 1$; $(R(1), R(2), R(3) \dots R(n))$ are the permutations of $(i = 1, 2, 3 \dots n)$ for which $S_{R(i-1)} \geq S_{R(i)}$ for all $i = 1, 2, 3 \dots n$.

In the following theorem, we prove that the aggregated values obtained from different DHLTSs by using the proposed ordered weighted aggregation operators are again DHLTSs. This means that the aggregation operators are valid. Additionally, we provide validation for the proposed ordered weighted aggregation operators using Yager–Dombi t-norms. □

Theorem 2. Suppose $S_i = \{F(S_{\alpha_i} \langle O_{P_i} \rangle) | i = 1, 2, 3 \dots n\}$ is a group of DHLTSs, then the aggregated value of DHLTSs using the DHLTYDOWA operator is also a DHLTS, that is:

$$DHLTYDOWA(S_1, S_2, \dots, S_i) = \bigoplus_{i=1}^n W_i S_{R(i)} = 1 - \frac{1}{1 + \left\{ \left(\sum_{i=1}^n W_i \right) \left\{ \frac{\min(1, F(S_{\alpha_i} \langle O_{P_i} \rangle))^t}{1 - \min(1, F(S_{\alpha_i} \langle O_{P_i} \rangle))^t} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}}} \tag{11}$$

Proof. The proof of this result is similar to the proof of Theorem 1. Moreover, the DHLTYDWA operator can easily hold its idempotence, boundedness, and monotonicity properties. Suppose $S_i^h = \{F(S_{\alpha_i} \langle O_{P_i} \rangle) | i = 1, 2, 3 \dots n\}$ is a collection of DHLTSs, then the double-hierarchy linguistic term Yager–Dombi hybrid weighted averaging (DHLTYDHW) operator is a mapping $\mathbb{Q}^n \rightarrow \mathbb{Q}$ such that:

$$DHLTYDHW A(S_1^h, S_2^h, S_3^h, S_3^h, \dots, S_i^h) = \bigoplus_{i=1}^n W_i S_i^h$$

$W_i = (W_1, W_2, W_3, W_4 \dots W_i)^t$ is the weight vector of $S_i^o = \{F(S_{\alpha i} \langle O_{P_i} \rangle) | i = 1, 2, 3 \dots n\}$ that fulfills the condition that $\sum_{i=1}^n W_i = 1$. \square

In the following theorem, we prove that the aggregated values obtained from different DHLTSs by using the proposed hybrid weighted aggregation operators are again DHLTSs. This means that the aggregation operators are valid. Additionally, we provide validation for the proposed ordered weighted aggregation operators using Yager–Dombi t-norms.

Theorem 3. Suppose $S_i^h = \{F(S_{\alpha i} \langle O_{P_i} \rangle) | i = 1, 2, 3 \dots n\}$ is a group of DHLTSs, then the aggregated value of DHLTSs using the DHLTYDHW A operator is also a DHLTS, that is:

$$DHLTYDHW A(S_1^h, S_2^h, \dots, S_i^h) = \bigoplus_{i=1}^n W_i S_i^h = 1 - \frac{1}{1 + \left\{ (\sum_{i=1}^n W_i) \left\{ \frac{\min(1, F(S_{\alpha i} \langle O_{P_i} \rangle)^t)}{1 - \min(1, F(S_{\alpha i} \langle O_{P_i} \rangle)^t)} \right\}^{\frac{k}{t}} \right\}^{\frac{1}{k}}} \tag{12}$$

Proof. The proof of this result is similar to the proof of Theorem 1. Moreover, the DHLTYDWA operator can easily hold its idempotent, bounded, and monotonicity properties. \square

4. Activation Functions and Neural Network Systems

An activation function [48] is a function that gives the output of a node. Additionally, it is known as the “transfer function.” It is used to determine if the output of a neural network is a yes or no response. Depending on the function, it can transfer the output values to a range between 0 and 1 or between −1 and 1, among others. An activation function can be either linear or nonlinear. The terms monotonic function and derivative are essential for understanding nonlinear functions. The range or curves of the nonlinear activation functions are the major factors used to categorize them. The sigmoid or logistic activation function that we use is defined as follows:

$$f(x) = Y = \frac{e^x}{e^x + 1} \tag{13}$$

where $x = 1 \pm 2 \pm 3 \pm 4 \dots \psi$.

The sigmoid activation function, which successfully accomplishes its duty and has a range of 0 to 1, is commonly employed in decision-making because it is effective at performing what it is supposed to perform. Since it has the lowest range and generates the most precise predictions, we therefore employ this activation function anytime we need to determine an outcome. The function could take many different shapes. As a consequence, we are able to determine the slope of the sigmoid curve between any two positions. Although the derivative of the function is not monotonic, the function itself is. The logistic sigmoid function has the effect of making it possible for a neural network to encounter an impasse during training. Neural networks were first developed in the 1950s as a means of addressing this issue. Programming everyday computers to function similarly to a network of brain cells is the process used to create ANNs. Artificial neural networks employ a complex mathematical algorithm to make sense of the data they are fed. A typical artificial neural network is made up of hundreds to millions of units, commonly referred to as artificial neurons, arranged in an order of layers. The input layer receives a variety of outside data types. The network’s goal is to process or comprehend these data. After leaving the input layer, the data go through one or more hidden units. The

hidden unit is in charge of transforming the incoming data into a format that the output unit can use. Most neural networks have complete interconnections between layers. These connections, similarly to the human brain, are weighted; the higher the number, the greater the impact one unit has on another. Every unit in the network gains knowledge as the data pass through it. The output units, which are placed on the other side of the network, provide the network with the data that have been received and processed. ANNs come in different types, but in this paper, we only discuss FFNNs in detail.

A feed-forward neural network (FFNN) is one in which there is no cycling of the connections between the nodes. The opposite of a FFNN is a feed-backward neural network (FBNN), which cycles through certain routes. The feed-forward model is the most fundamental type of neural network since incoming data are only ever processed in one way. In a feed-backward neural network, some of the input data come back to the input layer from hidden layers. However, regardless of the number of hidden nodes the data may pass through, they never flow backward and always move ahead. The input layer, hidden layer, and output layer are only a few of the layers that make up this collection of fundamental processing units. Each unit in the layer below it is linked to every other unit in the layer above it. Because they are not all made equally, each of these connections may have a different weight or strength. A feed-forward neural network is a function such that:

$$\beta = f\left(\sum_{j=1}^n W_j a_{ij}\right) \tag{14}$$

where the activation function is denoted by f ; a_{ij} is the input signal; W_j is the criteria weight; and β is the single output. The output of feed-forward neural network is shown in Figure 1.

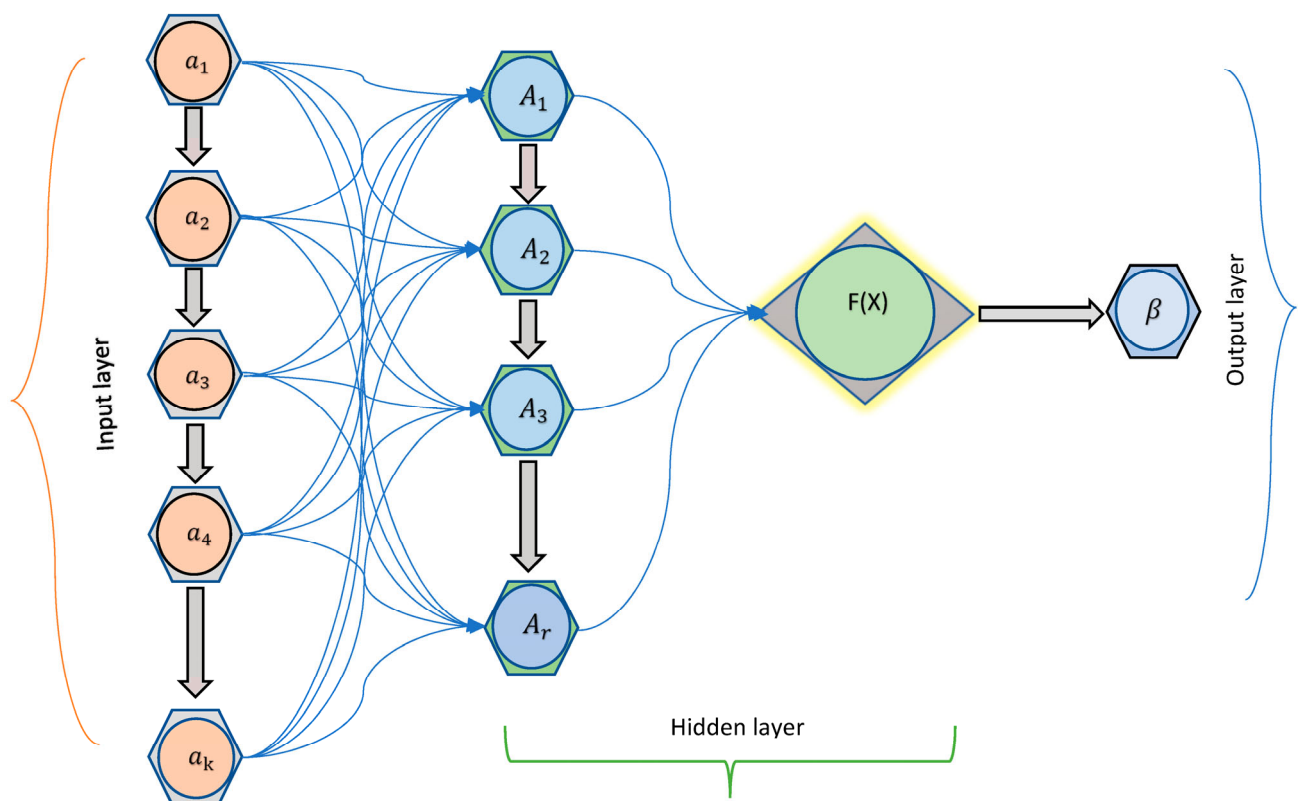


Figure 1. Output of a feed-forward neural network.

Figure 2 shows that the Artificial neural networks and fuzzy logic are both used in neuro-fuzzy systems, a sort of hybrid intelligent system that is effective at processing

complicated data. The following figure depicts a fuzzy neuron, a processing element of a hybrid neural net:

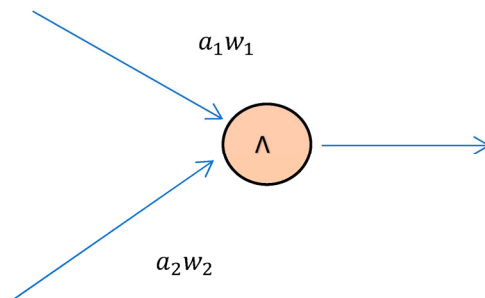


Figure 2. Simple neuro-fuzzy system.

where a_1, a_2 are the input signals and w_1, w_2 are the corresponding weights of the input signals, respectively. \wedge is the AND fuzzy neuron.

5. The Output of Neural Networks Using Yager–Dombi Operators

In this section, we put forward a new extension to the fuzzy neural system to develop a fuzzy neural system that is based on DHLTSs. For this, let $P_j = \{p_1, p_2, p_3, p_4 \dots p_k\}$ be the arrangement of attributes that will be evaluated, and let $Q_i = \{q_1, q_2, q_3, q_4 \dots q_z\}$ be the discrete set of z alternatives that must be selected. Suppose $W_j = (W_1, W_2, W_3, W_4 \dots W_k)^l$ is the attribute weight vector, where $W_j \{j = 1, 2, 3, 4 \dots k\}$ are all real numbers such that $W_j > 0$ and $\sum_{j=1}^z W_j = 1$. Let there be n number of experts, $DM_r \{r = 1, 2, 3, 4 \dots n\}$, to express their views on the z alternatives in reference to the k criterion in the context of DHLTSs. The data given by the n th expert in the form matrix appear as follows:

$$g_r = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{z1} & a_{z2} & \dots & a_{zk} \end{bmatrix}$$

To determine the output of the feed-forward fuzzy neural system, we go through different phases. In phase 1, we determine the criteria weight vector of each matrix provided by the experts corresponding to each input information or signal. In phase 2, we find the hidden layer using Equations (18) and (19). In phase 3, the activation function is used to calculate the output of the input data.

To determine the feed-forward fuzzy neural system’s output in a DHLTS environment, we suggest the following algorithm:

Phase 1: The following are the steps in Phase 1:

Step 1: Recognize the expert data or information presented in the form of a matrix $g_r = \{A_{ij}\}_{z \times k}$, where $j = (1, 2, 3, 4 \dots k)$ and $i = (1, 2, 3, 4 \dots z)$ represent the j th criterion value in relation to the i th alternative, respectively;

Step 2: Determine the criteria weight vector of each matrix provided by the experts corresponding to each input information or signal using the entropy measure method, which consists of the following steps:

1. To evaluate the entropy of the information provided by the experts in the form of matrices, we use the following equation:

$$n_j = \frac{1}{i} \sum_{i=1}^z \left(\frac{1}{j} \sum_{j=1}^k \left(\left\{ \sqrt{2} \cos \pi \left(\frac{F(S_{ai} \langle O_{Pi} \rangle) - (F(S_{ai} \langle O_{Pi} \rangle))^c}{4} \right) - 1 \right\} \times \frac{1}{\sqrt{2} - 1} \right) \right) \tag{15}$$

where $j = (1, 2, 3, \dots, k)$ and $i = (1, 2, 3, \dots, z)$ denote the j th alternative and i th criterium, respectively, and $F(S_{\alpha i} \langle O_{\mathbf{P}i} \rangle)$ denotes the DHLTS, $(F(S_{\alpha i} \langle O_{\mathbf{P}i} \rangle))^c$ is the complement of the given DHLTEs, and is determined as:

$$(F(S_{\alpha i} \langle O_{\mathbf{P}i} \rangle))^c = F(S_{-\alpha i} \langle O_{-\mathbf{P}i} \rangle) \tag{16}$$

2. We obtain the final criteria weight vector as follows:

$$W_j = \frac{n_j}{\sum_{j=1}^k n_j} \tag{17}$$

Phase 2: The steps in Phase 2 are listed below:

Step 1: Find the scalar product of the weight vector and input signal using the Yager–Dombi aggregation operator as follows:

$$\mathcal{M} = \sum_{r=1}^n (W_r a_r), \tag{18}$$

where $W_j = (W_1, W_2, W_3, \dots, W_k)$ and $r = (1, 2, 3, 4 \dots n)$ denote the weight vector and the experts, respectively. $a_j = (a_1, a_2, a_3, \dots, a_j)$ denotes the input signals;

Step 2: Take the minimum weight and multiply it with \mathcal{M} using the Yager–Dombi averaging operator.

$$W_j = \min\{\mathcal{G}_1(w_j), \mathcal{G}_2(w_j), \mathcal{G}_3(w_j), \dots, \mathcal{G}_n(w_j)\} \tag{19}$$

Phase 3: The steps taken in Phase 3 are as follows:

Step 1: Apply the logistic activation function to the outcome of Step 2 in Phase 2 to obtain the final output β . The logistic activation function is described as follows:

$$f(x) = Y = \frac{e^x}{e^x + 1}$$

Step 2: Rank the possible outcomes for each β in descending order.

Output of Feed-Forward Double-Hierarchy Linguistic Term Neural Networks

An artificial neural network of this sort, known as a feed-forward double-hierarchy linguistic term neural network (FFDHLTNN), only allows information to travel in one direction: from the input layer using one or more hidden layers to the output layer. It is called feed-forward because data move forward through the network without looping back on themselves. In a typical FFDHLTNN architecture, each layer consists of a set of neurons or nodes that perform a simple mathematical operation on the inputs and output the result to the next layer. The network’s ultimate output is produced by the output layer once the input layer has received the input data. As demonstrated in Figure 3, the hidden layers execute intricate modifications on the input data to provide characteristics that are beneficial for the output layer.

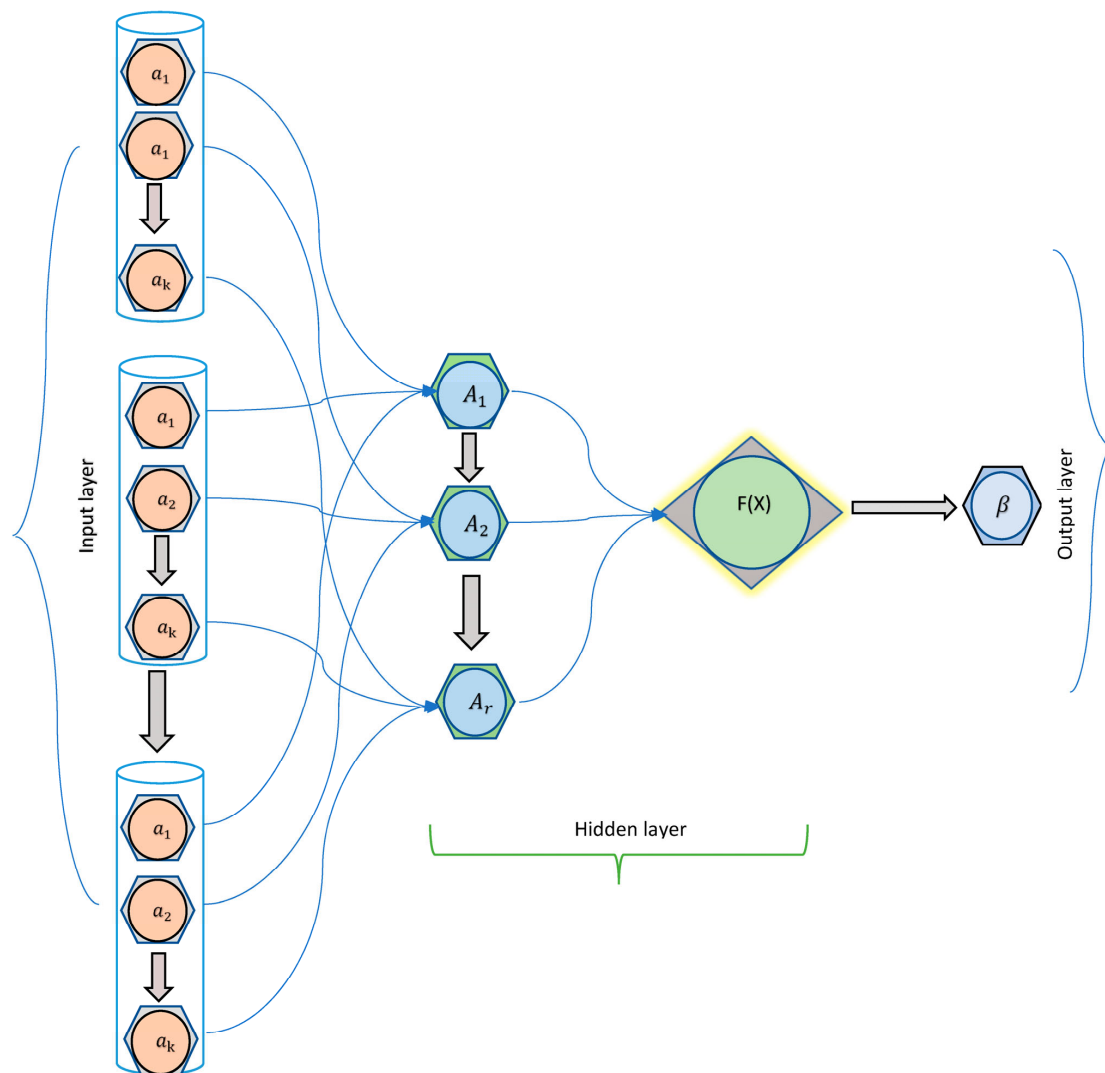


Figure 3. Output of a feed-forward double-hierarchy linguistic neural network.

Step 1: The decision-making matrix given by the decision-makers in the LTS environment must be identified, and all evaluations must be based on the LTS that the DM has provided:

$$S_o = (S_\alpha \langle O_P \rangle) \begin{bmatrix} S_{-4} = \textit{extremely low}, S_{-3} = \textit{very low}, S_{-2} = \textit{low}, S_{-1} = \textit{slightly low}, \\ S_0 = \textit{medium} \\ S_1 = \textit{slightly high}, S_2 = \textit{high}, S_3 = \textit{very high}, S_4 = \textit{extremely high} \\ O_{-4} = \textit{far from}, O_{-3} = \textit{scarcely}, O_{-2} = \textit{only a little}, O_{-1} = \textit{a little} \\ O_0 = \textit{just right} \\ O_1 = \textit{much}, O_2 = \textit{very much}, O_3 = \textit{extremely much}, O_4 = \textit{entierly} \end{bmatrix}$$

Step 2: Find the criteria weight for each matrix provided by the experts in the environment of the DHLTEs using the entropy measure method;

Step 3: Find the scalar product of input information or signals and their corresponding weights, and add the scalar product of the input signals using Equations (18) and (19);

Step 4: Apply the logistic activation function to the outcome of Step 2 in Phase 2 to obtain the final output β ;

Step 5: Rank the possible outcomes for each β in descending order.

6. Numerical Example

Water is a clear, flavorless, and odorless liquid that is necessary for all life on Earth. It is a chemical substance with the chemical formula H_2O , consisting of two hydrogen atoms and one oxygen atom. The three states of matter that water may exist in are solid (ice), liquid (water), and gas (water vapors). It is the only substance on Earth that can naturally exist in all three states at normal temperatures and pressures. Water plays a crucial role in many aspects of our lives, including hydration, agriculture, industry, transportation, and recreation. However, due to the growth of industries, access to clean and safe drinking water has become a serious concern for many people throughout the world. Water pollution causes various diseases, some of which lead to death. As a result, supplying clean water to domestic areas is the most crucial responsibility of governments. In this section, we will look at different commercial-scale water purification systems and the factors that influence them.

The following techniques are used to purify water on a commercial scale:

- (1) **q_1 : Chlorination:** Chlorination is a process that uses chlorine to disinfect water. Chlorine is added to water, which kills the harmful bacteria and viruses present in it. This method is effective in killing most of the disease-causing pathogens;
- (2) **q_2 : Reverse Osmosis:** In the reverse osmosis (RO) process, a semi-permeable membrane is utilized to filter out dissolved particles, contaminants, and minerals from water. It is a highly effective method of water purification and is commonly used in households and industries;
- (3) **q_3 : Ultraviolet Purification:** UV purification uses ultraviolet light to kill bacteria, viruses, and other microorganisms present in water. It is an effective method of water purification and does not use any chemicals;
- (4) **q_4 : Filtration:** Filtration is a process that removes impurities from water by passing it through a porous material. Sediments, dirt, and other bigger particles can be effectively removed from water with this technique;
- (5) **q_5 : Coagulation and sedimentation:** Chemicals such as alum are added to water to cause impurities to clump together and settle at the bottom of a tank, which can then be removed through sedimentation;
- (6) **q_6 : Boiling:** By bringing water to a boil for at least one minute, the majority of disease-causing organisms can be killed;
- (7) **q_7 : Distillation:** Water is heated during distillation, and the steam is subsequently condensed back into water. Minerals, chemicals, and bacteria are just a few of the impurities that can be removed by using this method.

The factors that affect the methods of water purification are:

1. **p_1 : Economic factors:** Economic factors can have a significant impact on water purification, as the process of treating and purifying water can be expensive and require significant investments in infrastructure, technology, and human resources. One major economic factor that can affect water purification is the availability and cost of resources such as energy, chemicals, and materials needed for the purification process;
2. **p_4 : Socio-political factors:** The socio-political environment can have a significant impact on water purification. Governments have an obligation to make sure that their populations have access to safe drinking water since access to clean water is a fundamental human right. However, the provision of clean water can be influenced by a variety of socio-political factors, such as social factors, public health, and political instability, among others;
3. **p_3 : Environmental factors:** There are many environmental factors that can affect water purification, including temperature, chemicals, turbidity, and climate change. Overall, environmental factors can have a significant impact on water purification and must be taken into account when designing and implementing water purification systems;
4. **p_5 : Type of Contaminants:** The type and concentration of contaminants present in the water will also affect the purification process. Different treatment processes are better

suited for removing different types of contaminants, such as chemicals, microbes, or sediments;

5. **p_5 : Water Quality Standards:** The level of purity required for the final product will affect the purification process. Different industries and applications have different standards for water quality, which will influence the choice of treatment process and the extent of purification required.

For this, let $(q_1, q_2, q_3, q_4, q_5, q_6, q_7)$ represent the set of alternatives for commercial-scale water purification. Let $(p_1, p_2, p_3, p_4, p_5)$ be the five variables influencing these procedures. The network’s ultimate output is produced by the output layer once the input layer has received the input data. Between the input and output layers, the hidden layers carry out complicated modifications to the data to provide characteristics that are beneficial to the output layer. The linguistic information is considered input signals, interacting with weight vectors in the input layer and using the Yager–Dombi t-norms to produce the product of input signals as a linguistic variable. Then, the aggregated information of input signals is calculated using Yager–Dombi t-conorms. The hidden layer information, after performing the Yager–Dombi t-norms and t-conorms, is given in Table 1. Apply the activation function to the hidden layer signals and determine the output layer signals of the linguistic feed-forward neural network. The output layer information is given in Table 2.

Table 1. Hidden layer of feed-forward neural network for DHLTSs.

	p_1	p_2	p_3	p_4	p_5
φ_1	0.738521429	0.326423333	0.212368533	0.136978263	0.1258662
φ_2	0.969678667	0.368916545	0.333768273	0.154811222	0.095050909
φ_3	0.570108667	0.487515667	0.185739733	0.104634095	0.087066391
\mathcal{M}	2.278308762	1.182855545	0.731876539	0.396423581	0.3079835
	0.694964668	0.541884482	0.422591636	0.283884909	0.235464362

Table 2. Output layer of feed-forward neural network for DHLTSs.

	q_1	q_2	q_3	q_4	q_5	q_6	q_7
DHLTYDWA	0.61904	0.66270	0.64047	0.63911	0.66773	0.63961	0.62284
DHLTYDOWA	0.62006	0.66972	0.64923	0.63942	0.67406	0.64974	0.62513
DHLTYDHWA	0.51688	0.52162	0.51856	0.51912	0.52149	0.51857	0.51747

According to experts, the best solution is reverse osmosis, whereas filtration is effective for basic water purification tasks, including chlorine and sediment removal. Reverse osmosis removes contaminants across a wider range. All pathogens in the water are eliminated by alternative methods. However, the dead bacteria still float in the water. On the other hand, an RO water purifier eliminates bacteria by filtering out their floating, dead corpses after they have been killed. As a result, RO-purified water is cleaner.

7. Verification of Our Proposed Method

In this section, we consider the extended TOPSIS approach and the GRA method to verify the effectiveness and validity of our proposed approach. The extended TOPSIS and GRA methods are used for verification in decision-making because they provide a more comprehensive and objective analysis of multiple criteria and are effective in handling uncertainty, imprecision, and complex decision-making scenarios. For this verification, we consider the information given by the experts in the form of three matrices in the environment of DHLTSs. The matrices provided by the experts consist of five attributes related to seven alternatives. We use the Yager–Dombi aggregation operator to aggregate the information given by the experts in the form of matrices in the environment of DHLTSs. Then, we apply the entropy approach to determine the criteria weight vector of the aggregated matrices. Finally, we apply the extended TOPSIS method to obtain the required output, which is given in Table 3.

Table 3. Output of extended TOPSIS method for DHLTSs.

	q_1	q_2	q_3	q_4	q_5	q_6	q_7
d_{ij}^+	0.21584	0.03857	0.14964	0.14933	0.04855	0.14578	0.22714
d_{ij}^-	0.07570	0.27066	0.1419	0.14222	0.24299	0.14576	0.0644
output	0.25967	0.87526	0.48672	0.48781	0.83347	0.49997	0.22089
raking	0.87526	0.83347	0.49997	0.48781	0.48672	0.25967	0.22089

Similarly, we use the Yager–Dombi aggregation operator to aggregate the experts’ information and then apply the entropy measure method to determine the criteria weight vector. Finally, we apply the GRA method to obtain the required output, as shown in Table 4.

Table 4. Output of GRA method for DHLTSs.

	q_1	q_2	q_3	q_4	q_5	q_6	q_7
d_{ij}^+	0.58842	0.87869	0.642025	0.70632	0.86793	0.66403	0.60736
d_{ij}^-	0.79424	0.52724	0.73448	0.70624	0.57306	0.677098	0.85416
output	0.42557	0.62499	0.46642	0.50003	0.60231	0.49513	0.41556
raking	0.62499	0.60231	0.50003	0.49513	0.46642	0.42557	0.41556

From Tables 2–4, we can see that the results obtained by using both methods (the extended TOPSIS method and the GRA method) are almost the same as the results obtained by using our proposed method. This ensures that our proposed method is valid. Reverse osmosis is therefore the ideal option for experts to choose, while filtering is appropriate for simple water jobs such as chlorine and sediment removal. Reverse osmosis removes contaminants across a wider range. All pathogens in the water are eliminated by alternative methods. However, the dead bacteria still float in the water. An RO water purifier, on the other hand, eliminates germs by killing them and filtering away their floating, lifeless bodies. Therefore, RO-purified water is cleaner.

8. Discussion and Comparison

In this section, we compare our suggested approach with the GRA method. For this, let $P_j = \{p_1, p_2, p_3, p_4 \dots p_k\}$ be the arrangement of attributes that will be evaluated, and let $Q_i = \{q_1, q_2, q_3, q_4 \dots q_z\}$ be the discrete set of z alternatives that must be selected. Suppose $W_j = (W_1, W_2, W_3, W_4 \dots W_k)^l$ is the attribute weight vector, where $W_j \{j = 1, 2, 3, 4 \dots k\}$ are all real numbers satisfying $W_j > 0$ and $\sum_{j=1}^z W_j = 1$. Let there be an n number of experts, $DM_r \{r = 1, 2, 3, 4 \dots n\}$, expressing their views on the z alternatives with respect to the k criterion in the context of DHLTSs. Here, the experts gave data in the form of three matrices in the environment of DHLTSs. The matrices provided by the experts consist of five attributes related to seven alternatives. First, we find the expert weight vector using the entropy measure method. We use the Yager–Dombi aggregation operator to aggregate the information provided by the experts in the form of matrices in the environment of DHLTSs and then apply the entropy measure method to calculate the criteria weight vector of the aggregated matrix, which is $W_j = (0.214625973, 0.201184071, 0.167857403, 0.204584654, 0.211747899)^l$. Next, we determine the aggregated matrix’s PIS and NIS. Finally, we apply the GRA method to determine the output and rank the possible outcomes of each output, as shown in Table 4. For this comparison, we will also use the extended TOPSIS approach. Both the extended TOPSIS and TOPSIS methods are MCDM techniques that seek to isolate the best alternative from a group of alternatives based on several criteria. The number of criteria that are taken into account is the fundamental distinction between TOPSIS and extended TOPSIS. Extended TOPSIS can accommodate many sets of criteria, whereas TOPSIS only takes into account one set of criteria. Similarly, in the extended TOPSIS method, we find the expert weight vector using the entropy measure method. We use the Yager–Dombi aggregation operator to aggregate the information provided by the experts in the form of

matrices in the environment of DHLTSs and then apply the entropy measure method to determine the criteria weight vector of the aggregated matrix. Next, we find the positive and negative ideal solutions of the aggregated matrix. Then, the expert information is converted to one set of criteria, and we finally apply the TOPSIS method to determine the output and rank the possible outcomes of each output, as shown in Table 3. The graphical representation of the comparison between the proposed method and other similar methods can be observed in Figure 4. The comparison of detail information is given in Table 5.

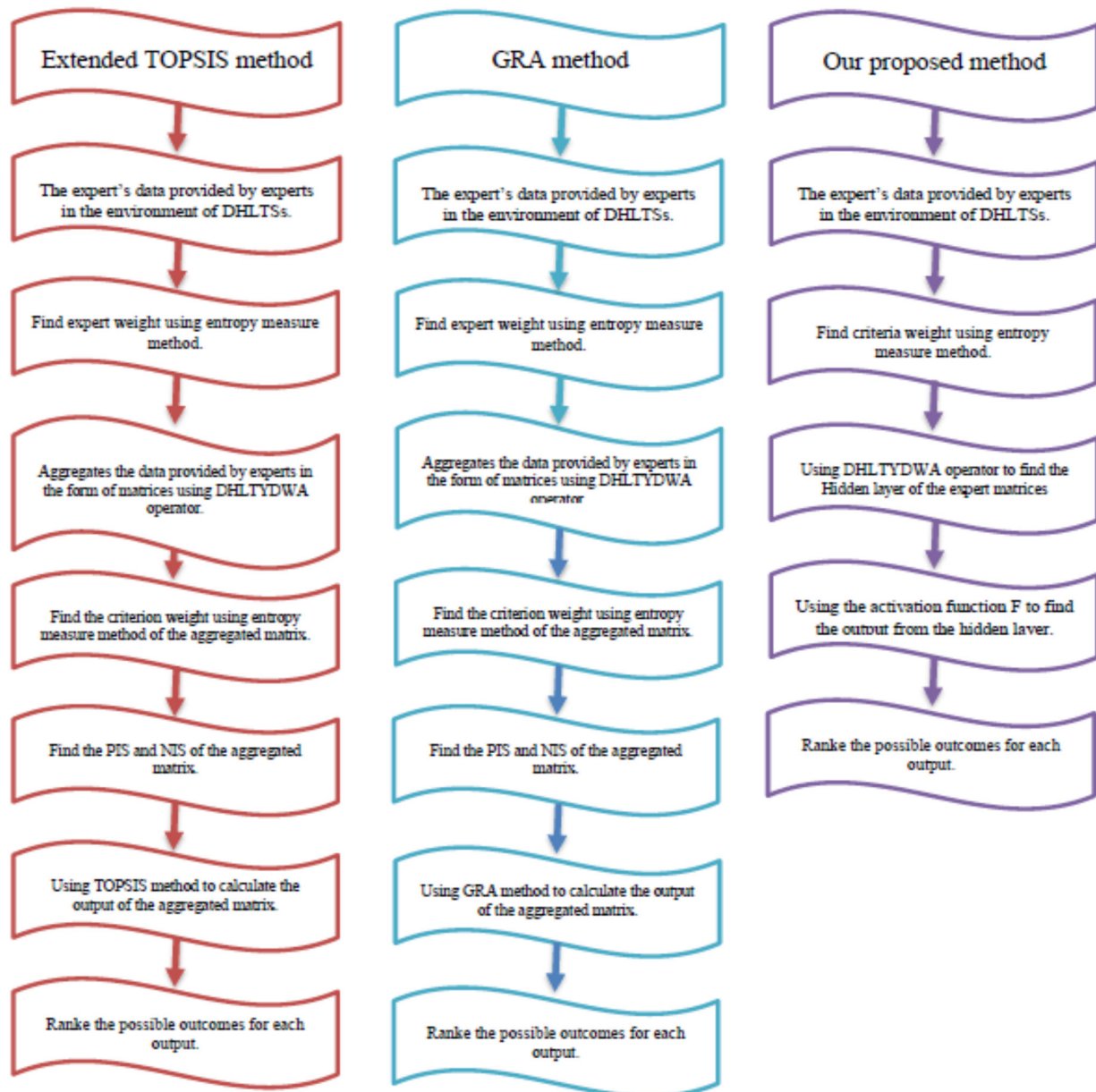


Figure 4. Steps of GRA method, extended TOPSIS method, and our proposed method.

Table 5. Output ranking of feed-forward neural network for DHLTSs.

DHLTYDWA	$q_5 > q_2 > q_3 > q_6 > q_4 > q_7 > q_1$
DHLTYDOWA	$q_5 > q_2 > q_6 > q_3 > q_4 > q_7 > q_1$
DHLTYDHWA	$q_2 > q_5 > q_4 > q_6 > q_3 > q_7 > q_1$

In our suggested approach, we take the experts’ information in the environment of DHLTSs and find the expert weight using the entropy measure method. After that, we use

the Yager–Dombi aggregation operator to find the hidden layer from the experts’ matrices, apply the activation function to generate the output of our proposed method, and finally rank the output. The comparison between the GRA method, the extended TOPSIS method, and our proposed method is shown in Figure 5. Both methods yield the same results as our proposed method. Filtration is useful for simple water tasks such as chlorine and sediment removal, but reverse osmosis is the best choice according to experts. A wider range of contaminants may be removed via reverse osmosis. Other techniques eliminate all germs in the water. However, the dead microorganisms still float in the water. An RO water purifier, on the other hand, filters out the dead bacteria that are floating in the water and eliminates them. It follows that RO-purified water is more sanitary. As a result, we conclude that our suggested technique is valid and superior to both methods (the extended TOPSIS method and the GRA method), since it produces the output in fewer steps, saving experts’ time.

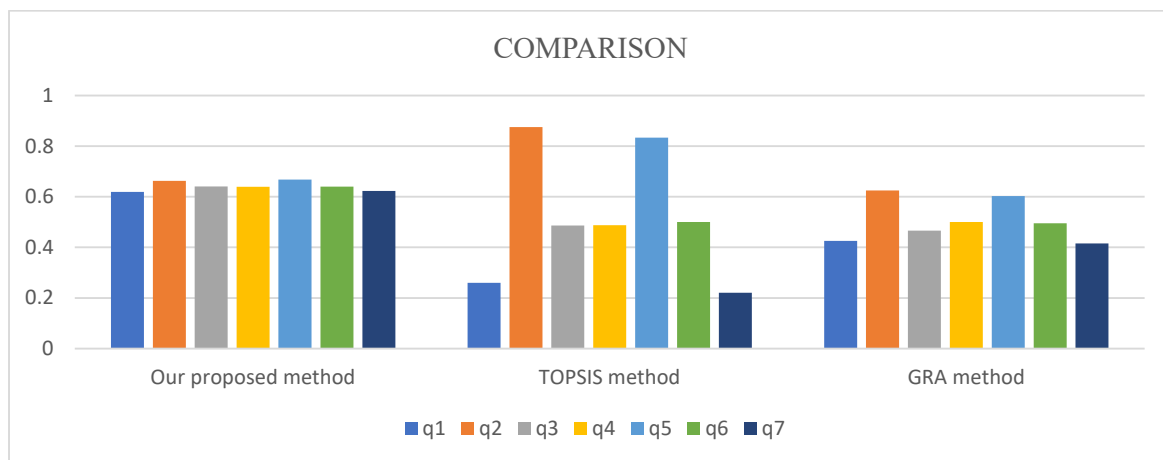


Figure 5. Comparison of Extended TOPSIS Method, GRA Method, and Our Proposed Method.

9. Conclusions

The neural network is a very important topic in machine learning, and artificial neural networks have garnered significant attention in the decision-making process. Linguistic information is a useful tool for describing uncertainty in information science and decision-making. Therefore, this paper applies the linguistic term set to artificial neural networks and develops a decision-making model based on linguistic neural networks. First, using the concepts of Yager t-norms and Dombi t-norms, we define a new hybrid t-norm known as the Yager–Dombi t-norm. We further develop some operational rules for double-hierarchy linguistic term sets and generalize them to incorporate more than two double-hierarchy linguistic term sets. The Yager–Dombi aggregation operator has been developed for double-hierarchy linguistic terms. We discuss the desirable properties of the DHLTYDWA operator, the DHLTYDOWA operator, and the DHLTYDHWA operator. Furthermore, we extend the concept of fuzzy neural network systems to feed-forward double-hierarchy linguistic neural network systems. Experts provide information in the environment of DHLTSs, and to determine the expert weight, we use the entropy measure method. We then obtain the hidden layer data using the DHLTYDWA operator and apply FFNNs to the hidden layer to derive the output of the information provided by the experts. In addition, a real-life MADM problem has been formulated. Filtration is beneficial for simple water tasks such as chlorine and sediment removal, but reverse osmosis is the best choice, according to experts, because a wider range of contaminants can be removed with this method. Other techniques eliminate all germs in the water. However, the dead microorganisms still float in it. An RO water purifier, on the other hand, filters out the dead bacteria that are floating in the water and eliminates them. It follows that RO-purified water is more sanitary. Moreover, we use the extended TOPSIS approach and the GRA approach for the verification of our proposed method, and both methods yield almost the same results as our proposed method.

A comparison analysis has been carried out, as seen in Figure 4, to demonstrate the validity and viability of our suggested method in comparison to other existing methods. We affirm that our proposed method is significantly better than other existing methods because it produces expert information output in a shorter time.

In future work, we will apply our proposed technique to intuitionistic fuzzy sets for decision-making problems and also use it in TWDs, Pythagorean fuzzy sets.

Author Contributions: Methodology, S.A. and N.A.; Formal analysis, A.O.A.; Data curation, A.O.A.; Writing—original draft, S.A. and N.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was funded by Institutional Fund Projects under grant no. IFPIP: 414-611-1443. The authors gratefully acknowledge the technical and financial support provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Data Availability Statement: Not applicable.

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

References

- Razi, M.A.; Athappilly, K. A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models. *Expert Syst. Appl.* **2005**, *29*, 65–74. [\[CrossRef\]](#)
- Coats, P.K.; Fant, L.F. Recognizing financial distress patterns using a neural network tool. *Financ. Manag.* **1993**, *1993*, 142–155. [\[CrossRef\]](#)
- Ceylan, H.; Bayrak, M.B.; Gopalakrishnan, K. Neural Networks Applications in Pavement Engineering: A Recent Survey. *Int. J. Pavement Res. Technol.* **2014**, *7*, 434–444.
- Sarvamangala, D.R.; Kulkarni, R.V. Convolutional neural networks in medical image understanding: A survey. *Evol. Intell.* **2022**, *15*, 1–22. [\[CrossRef\]](#)
- Fang, C.; Dong, H.; Zhang, T. Mathematical models of overparameterized neural networks. *Proc. IEEE* **2021**, *109*, 683–703. [\[CrossRef\]](#)
- Mitchell, J.M.O. Classical statistical methods. *Mach. Learn. Neural Stat. Classif.* **1994**, *1994*, 17–28.
- Koh, J.; Lee, J.; Yoon, S. Single-image deblurring with neural networks: A comparative survey. *Comput. Vis. Image Underst.* **2021**, *203*, 103134. [\[CrossRef\]](#)
- Alshehri, S.A. Neural network technique for image compression. *IET Image Process.* **2016**, *10*, 222–226. [\[CrossRef\]](#)
- Yen, Y.; Fanty, M.; Cole, R. Speech recognition using neural networks with forward-backward probability generated targets. In Proceedings of the 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, Munich, Germany, 21–24 April 1997; Volume 4, pp. 3241–3244.
- Collobert, R.; Weston, J. A unified architecture for natural language processing: Deep neural networks with multitask learning. In Proceedings of the 25th International Conference on Machine Learning, Helsinki, Finland, 5–9 July 2008; pp. 160–167.
- Ma, Q. Natural language processing with neural networks. In Proceedings of the Language Engineering Conference, Hyderabad, India, 13–15 December 2002; pp. 45–56.
- Rani, S.; Kumar, P. Deep learning based sentiment analysis using convolution neural network. *Arab. J. Sci. Eng.* **2019**, *44*, 3305–3314. [\[CrossRef\]](#)
- Chen, P.; Sun, Z.; Bing, L.; Yang, W. Recurrent attention network on memory for aspect sentiment analysis. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 7–11 September 2017; pp. 452–461.
- Draper, N.R.; Smith, H. *Applied Regression Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 1998; Volume 326.
- Yoo, P.D.; Kim, M.H.; Jan, T. Machine learning techniques and use of event information for stock market prediction: A survey and evaluation. In Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06), Sydney, Australia, 29 November–1 December 2006; Volume 2, pp. 835–841.
- Adya, M.; Collopy, F. How effective are neural networks at forecasting and prediction? A review and evaluation. *J. Forecast.* **1998**, *17*, 481–495. [\[CrossRef\]](#)
- Zaghloul, W.; Lee, S.M.; Trimi, S. Text classification: Neural networks vs support vector machines. *Ind. Manag. Data Syst.* **2009**, *109*, 708–717. [\[CrossRef\]](#)
- Naseer, M.; Minhas, M.F.; Khalid, F.; Hanif, M.A.; Hasan, O.; Shafique, M. Fannet: Formal analysis of noise tolerance, training bias and input sensitivity in neural networks. In Proceedings of the 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE), Grenoble, France, 9–13 March 2020; pp. 666–669.
- Shin, K.S.; Lee, T.S.; Kim, H.J. An application of support vector machines in bankruptcy prediction model. *Expert Syst. Appl.* **2005**, *28*, 127–135. [\[CrossRef\]](#)

20. Sterling, A.J.; Zavitsanou, S.; Ford, J.; Duarte, F. Selectivity in organocatalysis—From qualitative to quantitative predictive models. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2021**, *11*, e1518. [[CrossRef](#)]
21. Liu, Y.; Liu, S.; Wang, Y.; Lombardi, F.; Han, J. A survey of stochastic computing neural networks for machine learning applications. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 2809–2824. [[CrossRef](#)]
22. Schwendicke, F.A.; Samek, W.; Krois, J. Artificial intelligence in dentistry: Chances and challenges. *J. Dent. Res.* **2020**, *99*, 769–774. [[CrossRef](#)]
23. Ossowska, A.; Kusiak, A.; Świetlik, D. Artificial intelligence in dentistry—Narrative review. *Int. J. Environ. Res. Public Health* **2022**, *19*, 3449. [[CrossRef](#)]
24. Drakopoulos, G.; Giannoukou, I.; Mylonas, P.; Sioutas, S. The converging triangle of cultural content, cognitive science, and behavioral economics. In *Artificial Intelligence Applications and Innovations, Proceedings of the AIAI 2020 IFIP WG 12.5 International Workshops: MHDW 2020 and 5G-PINE 2020, Neos Marmaras, Greece, 5–7 June 2020, Proceedings 16*; Springer International Publishing: Berlin/Heidelberg, Germany, 2020; pp. 200–212.
25. Bebis, G.; Georgiopoulos, M. Feed-forward neural networks. *IEEE Potentials* **1994**, *13*, 27–31. [[CrossRef](#)]
26. Medsker, L.R.; Jain, L.C. Recurrent neural networks. *Des. Appl.* **2001**, *5*, 64–67.
27. Gu, J.; Wang, Z.; Kuen, J.; Ma, L.; Shahroudy, A.; Shuai, B.; Liu, T.; Wang, X.; Wang, G.; Cai, J.; et al. Recent advances in convolutional neural networks. *Pattern Recognit.* **2018**, *77*, 354–377. [[CrossRef](#)]
28. Lange, S.; Riedmiller, M. Deep auto-encoder neural networks in reinforcement learning. In Proceedings of the 2010 International Joint Conference on Neural Networks (IJCNN), Barcelona, Spain, 18–23 July 2010; pp. 1–8.
29. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [[CrossRef](#)]
30. Chao, J.; Shen, F.; Zhao, J. Forecasting exchange rate with deep belief networks. In Proceedings of the 2011 International Joint Conference on Neural Networks, San Jose, CA, USA, 31 July–5 August 2011; pp. 1259–1266.
31. Zavadskas, E.K.; Mardani, A.; Turskis, Z.; Jusoh, A.; Nor, K.M. Development of TOPSIS method to solve complicated decision-making problems—An overview on developments from 2000 to 2015. *Int. J. Inf. Technol. Decis. Mak.* **2016**, *15*, 645–682. [[CrossRef](#)]
32. Wei, G.W. GRA method for multiple attribute decision making with incomplete weight information in intuitionistic fuzzy setting. *Knowl.-Based Syst.* **2010**, *23*, 243–247. [[CrossRef](#)]
33. Guitouni, A.; Martel, J.M. Tentative guidelines to help choosing an appropriate MCDA method. *Eur. J. Oper. Res.* **1998**, *109*, 501–521. [[CrossRef](#)]
34. Elliott, D.L. *A Better Activation Function for Artificial Neural Networks*; Institute for Systems Research, Harvard University: College Park, MA, USA, 1993.
35. Schmidt-Hieber, J. Nonparametric Regression Using Deep Neural Networks with ReLU Activation Function. 2020. Available online: <https://arxiv.org/abs/1708.06633> (accessed on 30 May 2023).
36. Yin, X.; Goudriaan, J.A.N.; Lantinga, E.A.; Vos, J.A.N.; Spiertz, H.J. A flexible sigmoid function of determinate growth. *Ann. Bot.* **2003**, *91*, 361–371. [[CrossRef](#)] [[PubMed](#)]
37. Zamanlooy, B.; Mirhassani, M. Efficient VLSI implementation of neural networks with hyperbolic tangent activation function. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **2013**, *22*, 39–48. [[CrossRef](#)]
38. Kamruzzaman, J. Arctangent activation function to accelerate backpropagation learning. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.* **2002**, *85*, 2373–2376.
39. Montavon, G.; Samek, W.; Müller, K.R. Methods for interpreting and understanding deep neural networks. *Digit. Signal Process.* **2018**, *73*, 1–15. [[CrossRef](#)]
40. Sideris, A.; Orita, K. Structured learning in feedforward neural networks with application to robot trajectory control. In Proceedings of the 1991 IEEE International Joint Conference on Neural Networks, Singapore, 18–21 November 1991; pp. 1067–1072.
41. Sharma, S.; Mehra, R. Implications of pooling strategies in convolutional neural networks: A deep insight. *Found. Comput. Decis. Sci.* **2019**, *44*, 303–330. [[CrossRef](#)]
42. Garg, H.; Shahzadi, G.; Akram, M. Decision-making analysis based on Fermatean fuzzy Yager aggregation operators with application in COVID-19 testing facility. *Math. Probl. Eng.* **2020**, *2020*, 7279027. [[CrossRef](#)]
43. Akram, M.; Khan, A.; Borumand Saeid, A. Complex Pythagorean Dombi fuzzy operators using aggregation operators and their decision-making. *Expert Syst.* **2021**, *38*, e12626. [[CrossRef](#)]
44. Ye, F. An extended TOPSIS method with interval-valued intuitionistic fuzzy numbers for virtual enterprise partner selection. *Expert Syst. Appl.* **2010**, *37*, 7050–7055. [[CrossRef](#)]
45. Zadeh, L.A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets Syst.* **1978**, *1*, 3–28. [[CrossRef](#)]
46. Atanassov, K. Intuitionistic fuzzy sets. *Int. J. Bioautomation* **2016**, *20*, 1.

47. Li, X.; Xu, Z.; Wang, H. Three-way decisions based on some Hamacher aggregation operators under double hierarchy linguistic environment. *Int. J. Intell. Syst.* **2021**, *36*, 7731–7753. [[CrossRef](#)]
48. Cheng, F.; Liang, H.; Niu, B.; Zhao, N.; Zhao, X. Adaptive neural self-triggered bipartite secure control for nonlinear MASs subject to DoS attacks. *Inf. Sci.* **2023**, *631*, 256–270. [[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.