IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 3 February 2024, accepted 16 March 2024, date of publication 27 March 2024, date of current version 15 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3381627

RESEARCH ARTICLE

Accurate Agarwood Oil Quality Determination: A Breakthrough With Artificial Neural Networks and the Levenberg-Marquardt Algorithm

SITI MARIATUL HAZWA MOHD HUZIR[1], ANIS HAZIRAH 'IZZATI HASNU AL-HADI[2], ZAKIAH MOHD YUSOFF[1], (Member, IEEE), NURLAILA ISMAIL[2], (Member, IEEE), AND MOHD NASIR TAIB[2], (Senior Member, IEEE)

[1]Electrical Engineering Studies, College of Engineering, Universiti Teknologi MARA, Johor Branch, Pasir Gudang Campus, Masai, Johor 81750, Malaysia
[2]School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, Selangor 40450, Malaysia

Corresponding author: Zakiah Mohd Yusoff (zakiah9018@uitm.edu.my)

This work was supported by the Ministry of Higher Education under Fundamental Research under Grant FRGS/1/2023/ICT11/UITM/02/1.

ABSTRACT The agarwood oil quality has been divided into four grades, including low, medium-low, medium-high, and high, and has been thoroughly examined in this manuscript. Recently, there has been a high demand for agarwood oil but the current grading method is based on conventional techniques that rely on visual inspection of various characteristics such as intensity, smell, texture, and weight. However, this method is not standardized, making it difficult to grade agarwood oil accurately. Therefore, the use of artificial neural networks (ANN) in artificial intelligence (AI) was employed to develop a system for identifying agarwood oil quality using the Levenberg-Marquardt (LM) algorithm. Data from 660 samples of chemical compounds extracted from agarwood oil were used to train the ANN. To enhance the accuracy of agarwood oil quality identification with LM performance, the data was split into 70% for validation, 15% for training, and 15% for testing. The results showed that the ANN with the eleven inputs (10-epi-γ-eudesmol, α-agarofuran, γ-eudesmol, β-agarofuran, ar-curcumene, valerianol, β-dihydro agarofuran, α-guaiene, allo aromadendrene epoxide and Υ-cadinene) trained by ten hidden neurons of LM algorithm provided the best performance with 100% for accuracy, specificity, sensitivity and precision as well as minimum convergence epoch. The experimental implementation of the model was done using the MATLAB version R2015a platform. This study will help to standardize agarwood oil quality determination using intelligent modeling techniques and serve as a guide for future research in the essential oil industry.

INDEX TERMS *Aquilaria malaccencis*, ANN, Levenberg-Marquardt, multi-layer perceptron, agarwood oil.

I. INTRODUCTION

Over the past years, the link between agarwood oil quality and its price has been at the centre of much attention. The growing popularity of agarwood oil has led to its increased cost due to its specialty in medicinal properties to product medicines, therapeutic potential to product incense or any meditation purpose, and captivating fragrance to product perfumes [1], [2], [3]. However, the growing popularity of agarwood oil not aligned with its grading technique where it still using manual grading based on human sensory panel. This is based on industry based-problems.

It is believed that agarwood oil has the benefit of a calming effect on the mind and body, and can help reduce stress and anxiety. It has been used to treat conditions such as anxiety, depression, and post-traumatic stress disorder (PTSD) [4], [5]. In addition, in Southeast Asia, *Aquilaria malaccencis* species is the rarest and most expensive aromatic resinous

The associate editor coordinating the review of this manuscript and approving it for publication was Leimin Wang.

heartwood oil in the industry [6], [7]. Southeast Asian countries, particularly in Malaysia, Indonesia, Vietnam, and Myanmar, are renowned for their abundant and high-quality wild *Aquilaria* trees, which produce agarwood through a natural process triggered by either natural injuries or microbial deposits. These trees are characterized by their high resin content, making them a valuable resource for the production of agarwood oil and other related products [1], [8].

Previous published studies are limited by using human sensory panels only. This is why the agarwood oil grades also varies depending on the country, resulting in a diverse range of grading labels [9]. Each country has different marking in labelling the agarwood oil grades. For instance, in Japan, kanankoh and jinkoh represent the highest and lowest grades, respectively, while gaharu and kalambak are commercially available grades for agarwood in Malaysia [10], [11]. Similarly, in India, True agar, Bantang, Butha, and Dhum are the most popular agarwood grades [12]. Next in Vietnam and Papua New Guinea, the ABC term is used for agarwood grading, with group A representing the highest quality and quality decreasing for each subsequent letter. Researchers in Malaysia also use this ABC system to grade agarwood oil [13]. Additionally, high and low grade is the most basic grading system, with different aromas having varying values [4].

The main challenge faced by agarwood industry is the oil grades classification. Conventionally, the identification method of agarwood oil quality can be subjective and reproducible, it relies heavily on human expertise and visual appearances such as resin content, viscosity, weight, density, color intensity, and aroma [4], [14], [15]. However, the performance of agarwood classification is remains unclear and limited due to the simplistic approach of utilizing human assumptions [1], [16], [17]. There is currently no standard grading system for agarwood grades in the industry, which has been proven to be inefficient due to the fact that human inspiratory muscles can become fatigued when sniffing and dealing with multiple oil samples, taking a significant amount of time to obtain the desired result [18]. For example, in Malaysia, if agarwood oil sample has a darker color intensity compared to other samples, it automatically classifies as grade A (the highest grade) notice as good quality. Same goes to the aroma of agarwood oil samples, the strong the aroma, the highest the quality [13]. Hence, this existed technique that widely used in the agarwood industry should been overcome and broadened in recent years by using intelligent system.

Moreover, the classification of agarwood grades is crucial in determining its price. The price of agarwood varies based on individual preferences and application suitability [9]. Higher quality agarwood oils command higher prices due to increased demand and scarcity. The price of agarwood has risen significantly due to the near-extinction of agarwood-producing trees [19], [20]. Reforestation efforts have helped to ensure a sustainable supply of agarwood [21]. The global price of agarwood ranges from $20 (RM90.73) to $50,000 (RM226,825.00) per kilogram, with high-quality agarwood essential oil priced at up to $30,000 (RM136,095.00) to $50,000 (RM226,825.00) per liter [22], [23], [24]. The most valuable products derived from agarwood are wood pieces and essential oils [25]. Along with this growth in agarwood market, however, there is increasing concern over the technique to grade the agarwood oil since its quality is the marker to determine the agarwood price. In another words, when there is no standard classification of agarwood oil grades, then there is no consistent price to follow the standard test. It's all based on human expertise as a grader not based on a system.

In general, the study found a tendency to grade agarwood oil based on the application of the modeling system as a future change. Besides, the intelligent modeling technique that involves abundance percentage of chemical compounds as input was designed to mitigate the risk of sniffing fatigue or loss of motivation to explore the odors, thereby enabling the determination of grades. This innovative approach can effectively prevent price manipulation and ensure fairness in trading among sellers and buyers, protecting the interests of local and legal agarwood traders [25]. Moreover, based on a recently published manuscript, the chemical profile of agarwood oil is proposed to be used to model its quality [26], [27], [28]. Before creating the modelling system, the oil samples must go through the pre-processing technique to investigate the data. Researcher has identified several key chemical compounds, including 10-epi-γ-eudesmol, α-agarofuran, β-agarofuran, dihydrocollumellarin, γ-eudesmol, agarospirol, and aromadendrane, which are present in high-quality agarwood oil [26]. Another previous studies have also identified β-agarofuran, α-agarofuran, 10-epi-γ-eudesmol, γ-eudesmol, longifolol, hexadecanol, and eudesmol as important compounds for both high and low quality agarwood [27]. These compounds have been found to be significant in modeling the quality of agarwood oil [28]. It is now understood that chemical profiles play an important role in modeling the agarwood oil quality using intelligent system.

Furthermore, the current agarwood grading system can be overcome by leveraging the power of artificial intelligence (AI) to provide a more comprehensive and accurate assessment of agarwood oil quality. Table 1 summarized the conducted studies by several researchers and academics on distinguishing the quality of agarwood oil into two grades - high and low - using various modeling techniques, including self-organizing maps (SOM), photovoltaic methods, k-nearest neighbor algorithms (K-NN), thermogravimetric analysis (TGA), support vector machines (SVM), and artificial neural networks (ANN) [6], [23], [28], [29], [30], [31], [32], [33], [34], [35]. The resulting models have demonstrated an accuracy rate of 80% or higher in determining the quality of agarwood oil [28], [30]. Despite the progress made in this field, there are still areas that require further exploration and development. Many scholars have investigated the agarwood

oil quality using intelligent modeling, but the majority of studies have focused on two grades only and still not applied in the industry [30], [36]. By expanding the scope of the study to include four qualities; low, medium low, medium high, and high based on the significant chemical compounds present in agarwood oil, such as 10-epi-γ-eudesmol, α-eudesmol, α-agarofuran, and β-agarofuran, this study aims to fill this gap in the literature and contribute to the advancement of agarwood oil quality determination using intelligent modeling.

From the Table 1, SOM based grading technique with the implement of KNN was successfully validated the agarwood oil into high and low grades. Since SOM has been proved to be effective in grading the agarwood oil for two qualities, the researcher suggests to improve the quality classification into three grades; high, medium and low grades and see whether the modeling can prove the effectiveness for more grades [33]. There was a finding from another research on the SVM technique [35]. The model using Polynomial Kernel parameter with seven chemical profiles were used as input in for SVM. These SVM technique was found to get 100% of accuracy, sensitivity, specificity and precision in classifying only two agarwood oil grades [35]. Hence, future research proposes to upgrade the amount of chemical profiles and the input samples that exceed two grades. Besides, another important finding was that ANN model with the implemented of Scaled Conjugate Gradient (SCG) algorithm successfully classified three significant compounds into high and low quality [34]. Overall, the ANN model can implement as a classifier to grading the agarwood oil more than two grades with improvement of chemical profiles amount.

Moreover, there are several types of training algorithm that can be implement as prediction or classification system. However, this recent study only focus on one training algorithm. From the literature in Table 2, the best selected algorithm was identified to classify the agarwood oil grades into four qualities. What can be clearly seen in Table 2 is the comparison and difference between the training algorithms of ANN in modelling and prediction the given dataset from the previous published papers. The table delineates the performance metrics and outcomes associated with each training algorithm. A good training algorithm was observed for the model.

Another previous study developed an application using Artificial Neural Network (ANN) with Levenberg-Marquardt (LM) and Scale Conjugate (SCG) algorithm to model the effect of nanosilica concentration, length of the pipe, and fluid temperature on the fluid flow characteristic for pipeline mixture improvement [38]. The ANN model configurations were optimized using 1 to 20 hidden neurons. Both training algorithms were compared using mean square error (MSE) values. Both outputs for Levenberg-Marquardt and Scale conjugate trained ANN models had MSE values of 9.35×10^{-7} and 5.62×10^{-2}, respectively. However, Scale conjugate algorithm and Levenberg-Marquardt

TABLE 1. Agarwood grading system using artificial intelligence (AI) for only two grades.

AI Techniques	Review and summary	Ref.
Self-organizing maps (SOM)	Three chemical compounds, namely β-agarofuran, α-agarofuran, and 10-epi-φ-eudesmol, were used as inputs for Self-Organizing Maps (SOM). The datasets for training, testing and validation were then computed in order with ratio of 80 to 20. The silhouette values of each cluster were computed and examined for negative values in both training and testing. Only clusters with positive silhouette values were accepted. Results : The high-grade cluster (neuron 1) had better average silhouette values compared to the low-grade cluster (neuron 2) in both training and testing. The samples in the high-grade cluster were well matched to their own cluster, which was low, and poorly matched to their neighboring cluster.	[33]
k-Nearest Neighbor (KNN)	Ten chemical compounds; β-agarofuran, α-agarofuran, 10-epi-φ-eudesmol, φ-eudesmol, longifolol, oxo-agarospirol, hexadecanol and eudesmol, were used as inputs. The dataset was tested from k=1 to k=5. Results : The performance of the k-NN algorithm was measured, and the highest accuracy achieved by k-NN, which was above 83.3%, demonstrated that k-NN was a reliable classifier for grading the agarwood oil into high and low qualities.	[32]
Artificial neural network (ANN)	Three compounds as input which are β-agarofuran, α-agarofuran, and 10-epi-ϒ-eudesmol. During network training, the Scaled Conjugate Gradient (SCG) algorithm was used as the default classifier. The ANN architecture consisted of three layers, and the number of hidden neurons was varied between 1 and 10 in each hidden layer. Results : The SCG-ANN algorithm has successfully classified three significant compounds into high and low quality based on their default performance; MSE value which is 0.0789 makes the hidden neuron 3 as the best hidden neurons.	[34]
Support vector machines (SVM)	Seven chemical compounds, namely β-agarofuran, α agarofuran, 10-epi-ϒ-eudesmol, ϒ eudesmol, longifolol, hexadecanol and eudesmol, were used as inputs for SVM. Each row of the testing data is test on the model using Polynomial Kernel Parameter and classified the data either it is in group 1(low quality) or group 2 (high quality). Results : The SVM technique was found to be a reliable classifier for agarwood oil classification into two grades, as evidenced by the 100% of accuracy, sensitivity, specificity, and precision achieved in the classification behavior.	[35]

TABLE 1. *(Continued.)* **Agarwood grading system using artificial intelligence (AI) for only two grades.**

Thermogravimetric analysis (TGA)	The thermogravimetric analysis (TGA) and differential thermogravimetric analysis (DTGA) were employed as a quick, convenient, and precise technique for detecting high-boiler adulterants in essential oils using the most expensive agarwood oil. Results : The TGA pattern of pure samples reveals single-stage volatilization (agarwood oil: 110–260°C, castor oil: 340–500°C, coconut oil: 330–450°C and polyethylene glycol-400: 260–390°C), whereas the adulterated sample exhibits two-stage volatilization in both lower- and higher-temperature zone owing to the presence of agarwood oil and high-boiler adulterant.	[23]

algorithm got lower MSE values but more in favor to Levenberg-Marquardt algorithm where it more close to 0 compared with Scale conjugate algorithm. The study concluded that the Levenberg-Marquardt algorithm had the most significant influence on the fluid flow characteristic for pipeline mixture improvement with a small mean square error value [38]. Clearly, Levenberg-Marquardt (LM) was found to be suitable and can be used to identify the future Agarwood oil quality of four grades in corresponding areas with its abundance of chemical compounds data from *Aquilaria Malaccensis* species. This is because LM algorithm has been shown to be faster and more efficient than other algorithms in training neural networks.

In this recent study, it is shown how an artificial neural network (ANN) may be used to resolve these issues specialty for four grades. Among the major areas of machine learning, neural networks, such as Artificial Neural Networks (ANNs), have emerged as promising AI tools in this field. Compared to Support Vector Machines (SVMs), the performance of ANNs is highly dependent on the distribution of data between input and output variables with more than two classes [30]. Factors such as the selection of input variables, the architecture of the network, and the quantity of training, testing, and validating data have a significant impact on the accuracy of the model. ANNs are capable of performing tasks such as classification, clustering, prediction, forecasting, and pattern recognition [39], [40], [41]. This novel setup will investigate the use of Artificial Neural Networks (ANNs) for agarwood oil quality identification using the Levenberg-Marquardt (LM) algorithm.

The dataset was divided into training, testing, and validation sets with a ratio of 70:15:15. The significant compounds were trained separately and repeatedly, with the optimal number of hidden neurons ranging from 1 to 10. Hidden neurons 1 to 10 was chosen as an improvement from previous study on grading the agarwood oil into two grades using $k = 1$ until $k = 5$ demonstrated accuracy of 83.3% [32]. The best architecture of the developed ANN model was evaluated using

TABLE 2. **The performance of training algorithms existed in artificial neural networks (ANN).**

Modeling / Prediction	Summary and Performance of the Training Algorithms	Ref.
A mild steel turning operation	This research investigated the application of two optimization algorithms, Levenberg-Marquardt (LM) and Scaled Conjugate Gradient (SCG), in the training of artificial neural networks for predicting metal removal rate (MRR) and average surface roughness (Ra) during the turning of cylindrical mild steel. Results : Based on the graphical and statistical results, the Levenberg-Marquardt (LM) algorithm provided the best fit to the data, resulting in the lowest root mean squared error (RMSE) and highest coefficient of determination (R^2) values. LM algorithm achieved the best performance compared to SCG algorithm.	[37]
Model the effect of nanosilica concentration, length of the pipe, and fluid temperature on the fluid flow characteristic for pipeline mixture improvement	Hidden neuron 1 to 20 was used for model configurations. Levenberg-Marquardt (LM) and Scale Conjugate (SCG) were used as training algorithm in ANN. Both training algorithms were compared using mean square error (MSE) values. Results : The study concluded that the Levenberg-Marquardt algorithm had the most significant influence on the fluid flow characteristic for pipeline mixture improvement with a small mean square error value.	[38]
Carbon dioxide adsorption on carbon-based adsorbents modeling	The input used in the model was BET surface, mesopore volume, micropore volume, temperature, and pressure. The study using multilayer perceptron (MLP) and radial basis function (RBF) neural networks. The study was using multilayer perceptron (MLP) and radial basis function (RBF) neural networks. Results : The Multi-Layer Perceptron (MLP) deep neural network with Levenberg-Marquardt (LM) training algorithms performed significantly better than the Radial Basis Function (RBF) network in training a dataset of over 200 adsorbers, achieving a remarkable correlation coefficient of 0.9951.	[37]

metrics such as confusion matrices, accuracies, sensitivities, specificities, precisions, mean squared errors (MSEs), correlation coefficients (Rs), and the number of epochs [42], [43], [44]. The aim, contributions and methods of this study are outlined below.

1) Improve the agarwood grading system and establish a fair trade among sellers and buyers based on the exact agarwood oil quality, which requires further extensive machine learning techniques.
2) Offer an optimized deep-learning model on 660 datasets from 11 chemical compounds for determining the qualities of agarwood oil.
3) Utilize feedforward multilayer perceptron artificial neural network (ANN) with Lavenberg Marquardt

algorithm to sort agarwood oil chemical abundance and modeling four class qualities of low, medium low, medium high, and high grade based on data acquired from *Aquilaria*species compare to previous published research with only two class qualities.

4) The parameter adjustment for agarwood oil quality determination has increased the classification accuracy, demonstrating the effectiveness of the proposed approach.

The artificial intelligent technique, such as artificial neural network (ANN) is a computational model inspired by the structure and function of the human brain [31]. Thanks to their ability to learn and adapt to new situations without being explicitly programmed, they have been used in robotics, autonomous vehicles, natural language processing, decision making, and control systems [45]. The Levenberg-Marquardt algorithm can also be trained in ANN model to solve non-linear least squares problems, by first defining a set of equations that describe the relationship between the observed data and the predicted data [46], [47], [48]. These equations are typically non-linear, and they may involve multiple parameters that need to be estimated, and the goal is to find the best fit of the data to a mathematical model, which will be described as in equation (3).

A. ARTIFICIAL INTELLIGENT

Artificial neural networks (ANN) are considered supervised learning algorithms and are able to classify relationships between functions to achieve a desired level of accuracy. In the literature review, ANN are categorized into various types, including feed-forward (FF), Kohonen, radial-basis function (RBF), recurrent, and spiking neural networks [31], [49]. Previous research also shows that The Multi-Layer Perceptron (MLP) deep neural network with Levenberg-Marquardt (LM) training algorithms performed significantly better than the Radial Basis Function (RBF) network as achieving a remarkable correlation coefficient [37]. Feed forward network is a static network where produces only one way (direction) associated with given input and output values. FF network is the most suitable since the other network such as recurrent network lack on its system where frequently computes the output neuron every time the input pattern is presented. It is due to the feedback paths of the modified input [45]. The FF ANN model with a multilayer perceptron (MLP) has been mentioned by researchers in the field of machine learning to have several subsequent layers, including the input layer, output layer, and hidden layers. The first layer is the input layer, and the second layer is the hidden layer that contains an activation function to generate output by processing the data from the input layer. The output layer produces linear output as responses to the input layer. Every neuron in a layer is connected to each neuron in the layer above it. Generally, data is transferred from the input to the output layer in a forward manner or one-way direction [31], [50]. The FF MLP ANN basic working principle is depicted in Figure 1.

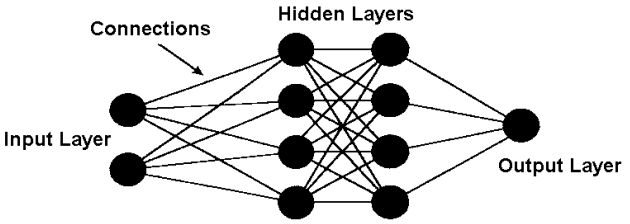

FIGURE 1. A multilayer perceptron architecture of ANN working principle.

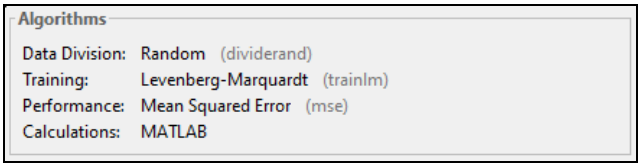

FIGURE 2. The LM algorithm used in ANN model using MATLAB software.

A previous non-linear ANN study on the identification of optimal grape maturity for harvest also provides better performance than linear models, such as Partial Least Square (PLS) [51]. The ANN model had an RMSE value of 156μg/kg and a correlation coefficient of 0.97, while the PLS model had an RMSE value of 867μg/kg and a correlation coefficient of 0.67. The study proved that the ANN model had the lowest mean square error and the highest correlation coefficient compared to the PLS model [51].

As a result, the ANN approach was found to be suitable and can be used to identify the future Agarwood oil quality in corresponding areas with its abundance of chemical compounds data from *Aquilaria Malaccensis* species. This is because Agarwood oil quality identification shares similarities with optimal grape maturity identification, where non-linear relationships exist between input and output data for classification systems [38], [51]. ANN (Artificial Neural Network) neural networks have the capacity to enable computers to make intelligent decisions with little human involvement. This ability arises from their ability to learn and represent complex systems with big data frameworks.

B. LEVENBERG-MARQUARDT ALGORITHM

The widely adopted method for training neural networks is the error backpropagation algorithm, also known as the steepest descent algorithm. Developed by Rumelhart and McCelland in 1986, this technique has been widely used in industry [52]. However, it has a relatively slow convergence rate, which has led to the development of enhancements to address this limitation. The slow convergence can be attributed to two main factors: large step sizes during optimization and the error surface's non-uniform curvature in all directions [53]. The Gauss-Newton algorithm was created to speed up the convergence of the steepest descent method. This algorithm is critical in determining appropriate step sizes for each direction and allows for relatively smooth convergence. However, its effectiveness is predicated on the

assumption that the error function's quadratic approximation remains reasonable throughout the optimization process [46], [47], [48].

The Levenberg-Marquardt (LM) algorithm has been proposed to overcome the limitations of the steepest descent method and Gauss-Newton method [54], [55]. The LM algorithm exhibits accelerated convergence in backpropagation, making it a widely preferred choice in practical applications and more robust than the Gauss-Newton method. LM has gained a reputation as an efficient and prominent algorithm when employed in an ANN system. It requires less training time while achieving superior regression accuracy [48].

The Hessian matrix and Jacobian matrix are two important matrices that are used in the Levenberg-Marquardt algorithm for optimization. The Levenberg-Marquardt algorithm was designed to achieve second-order training speed without requiring the computation of the Hessian matrix. Equation (1) expresses the relationship between the Hessian matrix (H) and the Jacobian matrix (J) where the Hessian matrix can be approximated, allowing for faster convergence and improved performance. Equation (2) expresses the calculation of gradient vector (g) [56]. There are defined as follows:

$$H = J^T \times J \tag{1}$$

$$g = J^T \times e \tag{2}$$

where the Jacobian matrix, J contains the first derivatives of the network errors with respect to the weights and biases, and the e is the vector of network error. The Jacobian matrix can be computed through a standard backpropagation technique, which is much less complex than computing the Hessian matrix.

Furthermore, equation (3) represents the modified equation as a result of the LM algorithm behaving as a new Hessian matrix into the Gauss-Newton method [40], which means that the LM algorithm uses the Jacobian matrix to approximate the Hessian matrix, which is scalar μ then used to compute the step size for each iteration of the algorithm. The scalar μ is reduced after each successful step (decrease in performance function) and is increased only when a tentative step would increase the performance function. This leads to a more efficient and stable optimization process, compute:

$$x_{k+1} = x_k - \left[J^T J + \mu I\right]^{-1} J^T e \tag{3}$$

where x_k indicate new weight calculated as gradient function, μ known as constant of dumping term and I is an identity matrix.

The Levenberg-Marquardt (LM) algorithm has been successfully applied to various identification tasks in ANN, including future paddy field identification under climate change, oil quality identification, and others [56], [57]. The paddy harvest prediction showed that the correlation coefficient (R) ranged between 0.65 and 0.8, indicating a strong positive correlation between paddy yield and climatic factors [56]. Although a perfect correlation coefficient of 1 would be ideal, the values of 0.65-0.8 are still considered acceptable.

Moreover, LM achieved the highest classification accuracy and outperformed the SCG algorithm for oil identification of two grades. At the lowest iterations and epochs, LM achieved the highest accuracy and lowest MSE value [57]. The correlation coefficient calculation (R) and mean square error value calculation (MSE) in the LM algorithm are represented by equations (4) and (5), respectively.

$$R = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{\left(N\sum x^2 - \left(\sum x\right)^2\right)\left(N\sum y^2 - (\sum y)^2\right)}} \tag{4}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2 \tag{5}$$

where x, y and N are the desired output value, the predicted output value and the number of data observations, respectively.

Therefore, the LM algorithm is a reliable and effective method for ANN-based identification tasks, offering a high degree of accuracy and stability. Its ability to handle complex data sets and non-linear relationships makes it a valuable tool for a wide range of applications.

The rest of this article is organized as follows. In Section II provides a detailed description of our proposed algorithm for model development. Our simulation results are shown in Section III. Finally, Section IV concludes this article.

II. METHODOLOGY

In this section, a comprehensive dataset is presented detailing the percentage abundance of chemical compounds, serving as input for the study. Subsequently, this meticulously collected dataset underwent training within the MATLAB software, utilizing an ANN training model for classification. In enhancing the ANN model's efficacy, one to ten hidden layers were incorporated in the architecture, with the dependent variable being the qualities of agarwood oil. The evaluation of the training algorithm's performance was conducted based on both the mean squared error (MSE) and correlation coefficient (R), ensuring a rigorous assessment of its effectiveness.

A. DATASET

Considering the manipulated variables essential for the experiment, the imperative need for collecting data on the abundance of agarwood chemical compounds becomes pivotal to ensure the accuracy of the grading for four distinct categories. The experimental analysis employs MATLAB software, utilizing a matrix-based language to handle a spectrum of tasks, from executing simple interactive commands to managing large-scale applications. The specific MATLAB version employed in this experiment is R2015a, chosen for its capability to construct, train, and evaluate the machine learning model focused on discriminating agarwood oil quality.

The agarwood oil dataset was obtained from FRM and BioAromatic Research Centre of Excellence (BARCE), Universiti Malaysia Pahang (UMP) as tabulate in Table 3 [28].

TABLE 3. Samples from FRIM and BARCE, UMP based on its quality [28].

Grades	Percentage	Number of Samples
High	330	50.00%
Medium High	30	4.50%
Medium Low	90	13.60%
Low	210	31.80%

A total of 660 samples, comprising 22 primary samples labeled as CKE, CM, EO2, EO3, EO4, HD, HG, JBD, KB, LA, LG, M, MA, MA1, MA2, MN, MNS, MPE, MS, R5, RG, and T, were utilized for this study. These samples, featuring 103 chemical compounds, underwent extraction and analysis through gas chromatography-mass spectrometry (GC-MS). The GC-MS apparatus settings included:

1) The initial temperature of the apparatus was set at 60 degrees Celsius and remained constant for 10 minutes.
2) The temperature gradually increased by 3 degrees Celsius per minute, ending at 230 degrees Celsius.
3) The flow rate of the helium gas carrier was set at 1 milliliter per minute.
4) The temperature of the ion source was set at 280 degrees Celsius.

By comparing the mass spectral patterns of the 660 samples to those in the mass spectral libraries (HPCH2205.L; Wiley7Nist05a.L; NIST05a.L), the eleven significant chemical compounds are identified with the help of a chemist.

The abundance parameters of these chemical compounds were integrated as input for the ANN model. These abundances, derived from eleven chemical compounds including 10-epi-γ-eudesmol, α-agarofuran, γ-eudesmol, β-agarofuran, ar-curcumene, valerianol, β-dihydro agarofuran, α-guaiene, allo aromadendrene epoxide, and Υ-cadinene. Simultaneously, the experiment's output involved identifying the quality of agarwood oil across different grades: high, medium-high, medium-low, and low grades.

B. MACHINE LEARNING CLASSIFICATION

The experimental step began with data pre-processing, which includes data normalization, randomization, and division. The data normalization technique used on the agarwood oil dataset was min-max scaling, which converts continuous input features into a range from 1 to 4 (high, medium-high, medium-low or low). After normalization, the rules for ANN classification were as follows:

1) The ratio of training, validation and testing: 70:15:15.
2) Number of hidden neuron: 10
3) Algorithm: Levenberg-Marquardt (LM)

The data was divided into three groups: training, validation, and testing, with a ratio of 70:15:15 for all data, respectively. Seventy percent of the data was allocated for training, leaving fifteen percent for validation to ensure the neural network is generalizing well and preventing overfitting. The remaining fifteen percent of the data was reserved for testing the network's performance.

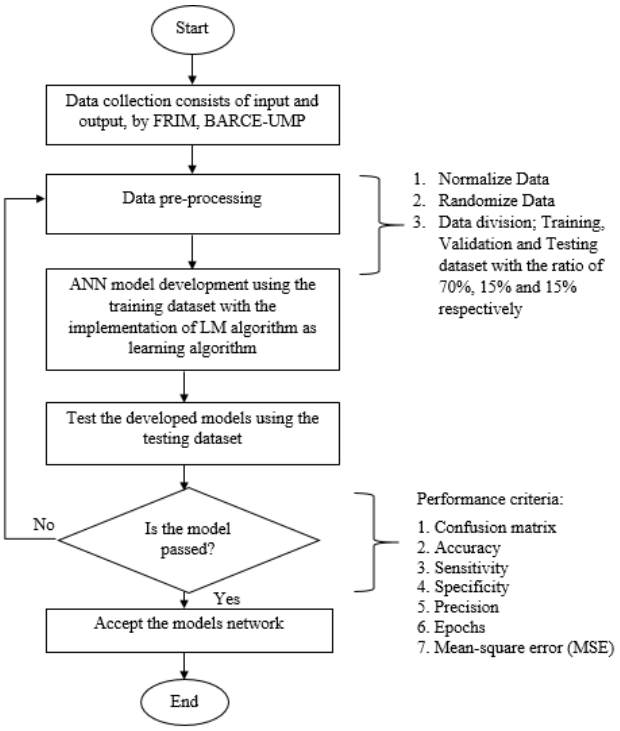

FIGURE 3. Process framework of artificial neural network (ANN) model.

Subsequent to data pre-processing, the dataset was trained and classified using an Artificial Neural Network (ANN) model. The chosen learning algorithm for this study was the Levenberg-Marquardt (LM) algorithm. Only one algorithm was implemented and evaluated in this study, as shown in Figure 3. The ANN architecture included one to ten hidden layers with the dependent variable of agarwood oil qualities. The performance of the training algorithm was evaluated based on the Mean Squared Error (MSE) and Correlation Coefficient (R) values. Mathematical expressions for R and MSE are provided as in Equations (4) and (5).

The established guidelines have been incorporated into the ANN model. In assessing the network's performance, additional metrics such as the confusion matrix and the number of epochs were examined. The evaluation of accuracy, sensitivity, specificity, and precision was conducted through the analysis of the confusion matrix output within the MATLAB numerical computing environment. The confusion matrix, a fundamental tool for evaluating model performance, encompasses four crucial metrics: true positives (TP), false negatives (FN), false positives (FP), and true negatives (TN). These metrics can be interpreted as follows:

1) True Positive (TP): This indicates that the model correctly identifies the oil quality as its grade. In other words, it correctly classifies as high grade oil quality when they are indeed high grade. Same goes to others 3 grades.

TABLE 4. The confusion matrix.

Data Group	Predicted (H)	Predicted (MH)	Predicted (ML)	Predicted (L)
Actual (H)	TP	FN_1	FN_2	FN_3
Actual (MH)	FP_1	TN_1	TN_2	TN_3
Actual (ML)	FP_2	TN_4	TN_5	TN_6
Actual (L)	FP_3	TN_7	TN_8	TN_9

'H' for High, 'MH' for Moderate-High, 'ML' for Moderate-Low, and 'L' for Low)

2) True Negative (TN_T): This represents that the model correctly identifies the Agarwood oil quality as not its grade. In simpler term, the agarwood oil is not high grade and the model also predicts the same which is not high grade.
3) False Positive (FP_T): Here, the model incorrectly predicts the oil quality to have not its grade. This is an erroneous classification of a not high grade (can be medium high, medium low or low) quality but classified as high grade, also known as a type-1 error.
4) False Negative (FN_T): In this case, the model wrongly predicts that the oil quality as its grade. i.e. high grade is incorrectly classified by the model as a not high grade (either medium high, medium low or low). This is known as a type-2 error.

In summary, the confusion matrix provides valuable insights into how well the model performs in distinguishing the agarwood oil quality of high, medium, medium low, and low grades. It identifies correct predictions (true positives and true negatives) and incorrect predictions (false positives and false negatives). All these metrics are derived from the confusion matrix outlined in Table 4.

In addition, a neural network that successfully achieves an optimal learning rate, resulting in rapid loss reduction within a limited number of epochs is thought to be capable of sustaining its performance on future test data without significant fluctuations in metrics such as accuracy and error. This suggests that the neural network is able to generalize well to new data and is not overfitting to the training data. The high number of neurons and epochs indicates that the relationship between the independent variables, specifically the significant agarwood chemical compounds and the dependent variable, was intricate and complex. This complexity is reflected in the fact that the neural network required a larger number of epochs to achieve optimal performance. However, the reduced processing time that comes with using fewer epochs makes it useful for handling and analyzing large datasets with large amounts of data.

Next, the accuracy of a model classifier will be the focus of attention, which reflects its overall effectiveness and can be interpreted as equation (6) below.

$$Accuracy = \frac{TP + TN_T}{TP + TN_T + FP_T + FN_T} * 100 \quad (6)$$

Sensitivity is a true positive rate that measures the model's ability to correctly predict positive cases of high quality group. It indicates that the model's prediction is positive and the Agarwood oil quality is high. The formula for calculating sensitivity is given as equation (7).

$$Sensitivity = \frac{TP}{TP + FN_T} * 100 \quad (7)$$

Specificity is a true negative rate that measures the model's ability to correctly predict negative cases of medium high, medium low, and low qualities group. The formula for calculating specificity is given as equation (8).

$$Specificity = \frac{TN_T}{TN_T + FP_T} * 100 \quad (8)$$

Precision is a metric that quantifies the proportion of true positive predictions made by the classification model or algorithm to the total number of positive predictions made. It is a measure of the accuracy of positive predictions. Precision can be calculated using equation (9).

$$Precision = \frac{TP}{TP + FP_T} * 100 \quad (9)$$

The error of MSE is calculated using equation (5). Additionally, the model is accepted if the neural network of machine learning passes all the criteria mentioned above, as shown in Figure 3.

Finally, the ANN based Agarwood oil quality grading technique was validated using independent samples to pass all the criteria explained above in term of its confusion matrix, accuracy, sensitivity, specificity, precision, number of epochs and MSE value using MATLAB software.

III. EXPERIMENTAL RESULTS

Table 5 presents a comprehensive view of the chemical profiles selected through the pre-processing technique of Principal Component Analysis (PCA). The PCA method successfully identified 11 significant chemical compounds, which were subsequently utilized as input for ANN modeling. Each chemical compound is listed in the table along with its respective values along the first two principal components, denoted as r-PC1 and r-PC2.

The chemical compounds included in the table are α-guaiene, β-agarofuran, ar-curcumene, β-dihydro agarofuran, γ-cadinene, α-agarofuran, 10-epi-γ-eudesmol, γ-eudesmol, allo aromadendrene epoxide, Valerianol, and Dihydrocollumellarin. The values of r-PC1 and r-PC2 provide insights into the distribution and variability of each chemical compound in the dataset. These values are essential for understanding the relative importance and contribution of each compound to the overall dataset, as determined by the PCA technique. The positive and negative values in the r-PC1 and r-PC2 columns signify the direction and magnitude of the contributions of each chemical compound to the identified principal components. For example, positive values indicate a positive correlation, while negative values indicate a negative correlation. These 11 identified chemical compounds and

TABLE 5. Eleven important chemical compounds that used as input for ANN modeling [28].

Name of Chemical Compounds	r-PC1	r-PC2
α-guaiene	0.059	0.613
β-agarofuran	0.673	0.030
ar-curcumene	-0.307	0.618
β-dihydro agarofuran	0.303	0.671
γ-cadinene	-0.798	0.088
α-agarofuran	0.814	0.186
10-epi-γ-eudesmol	0.667	0.158
γ-eudesmol	-0.214	0.756
allo aromadendrene epoxide	-0.246	0.641
Valerianol	0.731	-0.132
Dihydrocollumellarin	0.053	0.760

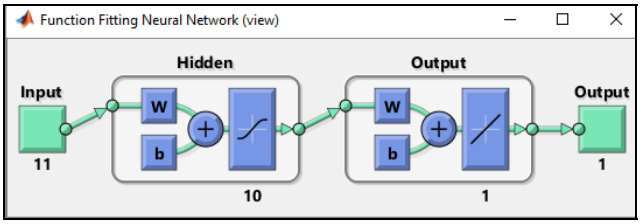

FIGURE 4. The neural network with 10 hidden layer.

```
%% load input-output data to the workspace
load('D:\Hazwa\Downloads\PHD\Research\DataAllGrades.mat');
load('D:\Hazwa\Downloads\PHD\Research\Grade4.mat');

%set initial weights
RandStream.setGlobalStream ...
     (RandStream ('mtl9937ar','seed',1));

x = DataAllGrades(:,1:11);
y = Grade4;
```

FIGURE 5. Input and output data.

```
net = patternnet(10,'trainlm','mse');
[net,tr] = train(net,x',y');
[trainInd, valInd, testInd] = dividerand (660, 0.7, 0.15, 0.15);
```

FIGURE 6. LM algorithm trained with the division of 70:15:15 for training, validation, and testing.

their associated principal component values serve as valuable input features for the ANN modeling process. The use of PCA in the pre-processing phase not only aids in dimensionality reduction but also ensures that the selected chemical profiles carry the most pertinent information for the subsequent modeling steps. The incorporation of these compounds into the ANN model is expected to enhance the model's ability to capture and analyze the intricate relationships within the chemical data, ultimately contributing to the robustness and accuracy of the model.

Figure 4 shows the variation of hidden neurons ranging from 1 to 10 in a single hidden layer, while the output layer includes one neuron representing oil quality, which is classified as high, medium high, medium low, or low. Each hidden neuron is a node that can process and transform input data, adding to the neural network's overall complexity and representational capacity. This hierarchical structure allows the ANN model to effectively learn and distinguish intricate patterns within the input data, facilitating accurate classification and prediction of agarwood oil quality based on its chemical composition and associated attributes.

The dataset, denoted as ".mat" files, encapsulates the original data extracted from CG-MS analysis, identifying the chemical compounds present in agarwood oil. Specifically, the input data encompasses the abundance, represented as a percentage of peak area, of significant chemical compounds. Concurrently, the output is expressed in binary form: "1000" denotes high quality, "0100" signifies medium-high quality, "0010" represents medium-low quality, and "0001" indicates low quality. The file is named "DataAllGrades.mat," and the associated Matlab script, featured in Figure 5, facilitates the loading of input-output relationships. This figure illustrates the Matlab code employed for declaring the input-output configuration during the training and testing commands, providing a visual representation of the data processing flow.

Subsequently, the learning algorithm derived from the Levenberg-Marquardt method is graphically represented in Figure 6, highlighting the specific dataset referenced in the script.

Figures 7(a)-7(d) visually depict the correlation coefficient associated with the eleven chemical compounds concerning agarwood oil qualities, utilizing four grades as distinct data points, each corresponding to the LM training algorithm implementation. The outcomes of this analysis affirm that the ANN trained with the Levenberg-Marquard algorithm demonstrates a remarkable ability to predict agarwood oil identification with a high degree of accuracy. Across all four analytical scenarios: training, validation, testing, and overall, a robust correlation between predicted and actual values is evident, consistently yielding correlation coefficient values equal to one.

This signifies a perfect correlation, indicating that the independent variables effectively account for 100% of the variations in the dependent variables. As articulated in the methodology section, a correlation coefficient of 1 is considered optimal for accurate prediction. Furthermore, the tabulated R values elucidate that the predictive performance improves with an increased number of data points, underscoring the reliability and effectiveness of the model in capturing the intricate relationships within the dataset. Notably, the correlation coefficient values have consistently registered at unity, indicating perfect alignment between the predicted and

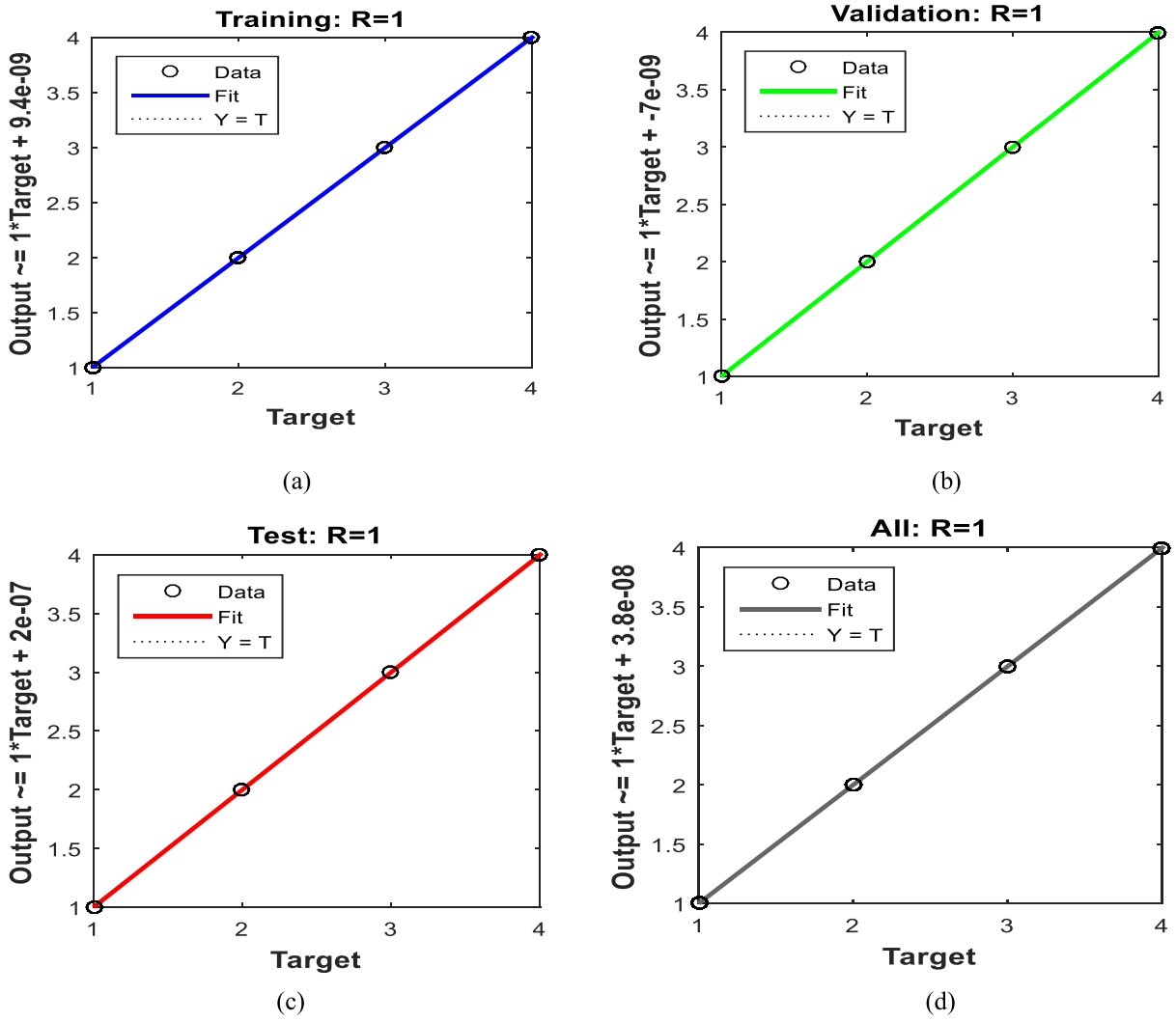

FIGURE 7. Coefficient correlation (R) for (a) training, (b) validation, (c) testing and (d) all under LM algorithm.

actual outcomes. This observation demonstrates the model's ability to accurately capture the intricate relationship between chemical compositions and agarwood oil qualities, allowing for precise classification and grading across multiple grades. The results demonstrate the efficacy of the LM training algorithm in optimising model convergence and improving predictive performance, reinforcing its utility in developing reliable and robust predictive models for agarwood oil analysis. As a result, the ANN model does not function effectively when data is scarce. This provides a clear justification for conducting the analysis annually by combining data from different Agarwood grades as four separate datasets.

Figures 8(a)-8(c) visually present the outcomes derived from the Levenberg-Marquardt (LM) algorithm's confusion matrix, meticulously evaluating the training, validation, and testing datasets. The allocation of data in a 70:15:15 ratio for the ANN model facilitated an impeccably harmonized correspondence between the predicted and actual data, attaining a flawless performance without any misclassifications during the testing phase. The confusion matrix vividly underscores the remarkable proficiency of the ANN model across four distinct grades: high, medium-high, medium-low, and low. In the testing phase, data was categorized into classes, with '1' denoting high grade, '2' representing medium-high grade, '3' indicating medium-low grade, and '4' signifying low grade. The meticulous evaluation of performance criteria for the ANN model revealed an outstanding 100% accuracy, sensitivity, specificity, and precision, attesting to its impeccable predictive capabilities.

Table 6 outlines a comprehensive summary of sample distribution across four grades within a dataset, categorized into training, validation, and testing sets. In the training set, constituting 70% of the dataset, there are 225 high-grade samples (48.70%), 21 medium-high samples (4.50%), 67 medium-low samples (14.50%), and 149 low-grade samples (32.20%). The validation set, comprising 15% of the dataset, consists of 48 high-grade samples (48.50%), 4 medium-high samples (4.00%), 16 medium-low samples (16.20%), and

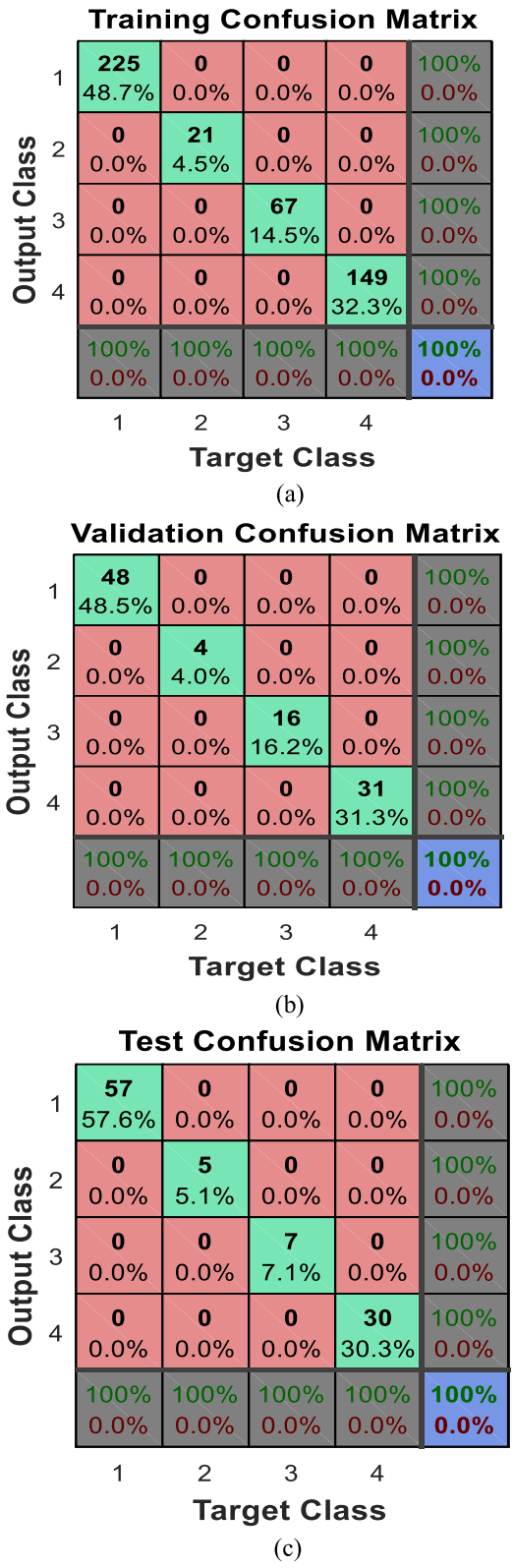

FIGURE 8. Confusion matrix of four different Agarwood grades of LM algorithm for (a) training, (b) validation and (c) testing dataset.

31 low-grade samples (31.30%). Lastly, the testing set, also representing 15% of the dataset, includes 57 high-grade samples (57.60%), 5 medium-high samples (5.10%), 7 medium-low samples (7.10%), and 30 low-grade samples (30.30%).

TABLE 6. Summary of samples number for four grades.

Dataset	Grades	Number of Samples	Percentage
Training (70%)	High	225	48.70%
	Medium High	21	4.50%
	Medium Low	67	14.50%
	Low	149	32.20%
Validation (15%)	High	48	48.50%
	Medium High	4	4.00%
	Medium Low	16	16.20%
	Low	31	31.30%
Testing (15%)	High	57	57.60%
	Medium High	5	5.10%
	Medium Low	7	7.10%
	Low	30	30.30%

TABLE 7. Performance criteria of four grades.

Accuracy	Sensitivity	Specificity	Precision
100%	100%	100%	100%

Table 7 present the performance criteria for four grades reveals that the LM training algorithm has achieved remarkable results in identifying relationships between chemical compound abundance and Agarwood oil quality determination. With a consistent 100% accuracy, sensitivity, specificity, and precision across all grades, the algorithm demonstrates exceptional reliability and effectiveness. This means that the model accurately classified chemical compounds in relation to agarwood oil quality, showing a high level of precision in its predictions. The perfect sensitivity indicates the algorithm's success in capturing all instances of relevant compounds associated with high-quality agarwood oil. Overall, these results signify the LM training algorithm's robust performance, laying a solid foundation for its practical application in determining Agarwood oil quality based on chemical compound abundance.

Table 8 presents a comprehensive evaluation of an ANN model based on various metrics, including accuracy, Mean Squared Error (MSE) value, and the number of epochs. The table outlines the performance of the model for different configurations of hidden neurons. The table illustrates the ANN model's accuracy across different hidden neuron configurations. Strikingly, the accuracy is consistently 100% for all configurations, including training, validation, and testing sets. This suggests that the model performs exceptionally well in correctly classifying instances, indicating a robust and reliable predictive capability. In this context, the MSE values are extremely low for all configurations, ranging from 0.048772×10^{-08} to 1988.0×10^{-08}. The low MSE values indicate that the model's predictions closely align with the actual values, reflecting high precision and minimal error in its predictions. The number of epochs represents the number

TABLE 8. The ANN model's evaluation based on accuracy, MSE Value and number of epoch.

Hidden Neuron	Accuracy (%)			MSE Value	Epoch
	Training	Validation	Testing		
1	100	100	100	1988.0×10^{-08}	31
2	100	100	100	33.073×10^{-08}	14
*3	100	100	100	0.048772×10^{-08}	22
4	100	100	100	1.7306×10^{-08}	13
5	100	100	100	1.8773×10^{-08}	12
6	100	100	100	1.4681×10^{-08}	11
7	100	100	100	0.26809×10^{-08}	12
8	100	100	100	0.32082×10^{-08}	12
9	100	100	100	0.23494×10^{-08}	11
10	100	100	100	1.5379×10^{-08}	10

*Best hidden neuron in LM algorithm

of times the algorithm iterates over the entire training dataset. The table provides the number of epochs corresponding to each hidden neuron configuration. It is notable that the number of epochs varies, with values ranging from 10 to 31. The lower number of epochs suggests that the model converges quickly and efficiently, optimizing its performance with fewer iterations.

The consistent 100% accuracy across all datasets, coupled with very low MSE values, indicates the ANN model's strong predictive power and generalization to unseen data. The variation in the number of epochs suggests that different hidden neuron configurations may require different training durations, with some configurations converging faster than others. The notably low MSE values underscore the model's ability to minimize prediction errors effectively.

While the study yields promising results for grading agarwood oil with an Artificial Neural Network (ANN), several limitations should be considered. To begin, the study focuses primarily on the application of ANN models to *Aquilaria*species, which may limit the findings' applicability to other agarwood variants. Extending the analysis to include a broader range of agarwood species may provide a more complete picture of the model's performance across different samples. Furthermore, relying on a single computational approach, namely ANN, may result in overlooking alternative methodologies that could provide complementary insights into agarwood grading. Exploring hybrid models or combining multiple machine learning techniques may improve the robustness and reliability of grading systems. Furthermore, the study's focus on computational efficiency and model performance metrics may have overlooked qualitative aspects of agarwood oil grading, such as aroma complexity and sensory attributes, which are important in market valuation and consumer preferences. Incorporating sensory evaluation methods and expert assessments could provide a more comprehensive view of agarwood quality. Furthermore, the study's use of a predetermined dataset and experimental setup may overlook the dynamic nature of agarwood production and market trends, necessitating ongoing validation and adaptation of grading models to meet changing industry requirements. Finally, data quality, sample representativeness, and potential biases must be carefully considered to ensure the reliability and validity of grading results. Addressing these limitations through interdisciplinary collaborations and methodological refinements can help to improve the applicability and robustness of agarwood oil grading methodologies in real-world settings.

IV. CONCLUSION

The FF MLP ANN were utilized to reveal the intricate nonlinear relationships among 11 chemical compounds and four grades of agarwood oil quality. A comparative analysis employing the Levenberg-Marquardt (LM) training algorithm demonstrated the LM algorithm's effectiveness, producing satisfactory and accurate results in the ANN. The findings revealed that the ANN performance had the highest accuracy by modeling using the LM algorithm. As a consequence, the LM algorithm presents a viable option for future predictions with reduced computational costs. The correlation coefficients obtained from the tests were optimal for predictions, precisely reaching a numerical value of 1. Moreover, as supported by existing literature, these correlation coefficients are considered acceptable, given their capacity to significantly enhance classification accuracy through parameter adjustments. In addition, ANN model trained by ten hidden neurons of LM algorithm provided the best performance with 100% for accuracy, specificity, sensitivity and precision as well as minimum convergence epoch. Thus, it can be confidently concluded that there exists a discernible correlation between agarwood oil quality and the abundance of chemical compounds from *Aquilaria* species. These findings underscore the efficacy of ANN methodologies in accurately grading agarwood oil, offering valuable insights into the domain of computational intelligence in natural product analysis. The novelty of this study also revealed that ANN system with the implemented of Levenberg training

algorithm can be applied to grade 4 qualities of agarwood oil; low, medium low, medium high and high grades for *Aquilaria* species.

Moreover, this prediction holds significant importance for consumers and traders, as it allows for AI-based determination on the oil prices based on agarwood oil qualities in the future. Intelligent modern techniques provide near-perfect predictions in contrast to manual grading which can be time-consuming and energy-intensive. In addition, consumers and traders anticipate acceptable identifications from these advanced AI methods. However, for the specific goals of our study, the chosen dataset provided valuable insights into the model's predictive capabilities within the specified grading scale. It would be interesting if the ANN model or others machine learning such as SVM, KNN, SOM and many more to be re-executed and achieve more accurate grading of Agarwood oil for various grades such as grades 5 and 6. Future research in the field of grading agarwood oil using Artificial Intelligence (AI) can focus on broadening the scope beyond *Aquilaria* species to include a wider range of agarwood variants. Researchers can gain a thorough understanding of the complexities involved in agarwood oil classification by incorporating a variety of agarwood species, including *Gyrinops*, *Gonystylus*, and *Aetoxylon,* into the grading process. This expansion enables the creation of AI models capable of accommodating the distinct chemical compositions and aromatic profiles found in various agarwood species, improving the accuracy and robustness of grading systems.

ACKNOWLEDGMENT

The authors would like to thank Universiti Teknologi MARA (UiTM), Johor Branch, and the School of Electrical Engineering, UiTM Shah Alam, Selangor, as excellent supportive institutes.

REFERENCES

[1] A. N. M. Ramli, S. Yusof, P. Bhuyar, A. W. Aminan, and S. N. Tajuddin, "Fungi mediated agarwood (A. Malaccensis) production and their pharmaceutical applications: A systematic review," *Int. J. Plant Based Pharmaceuticals*, vol. 2, no. 2, pp. 261–270, 2022.

[2] R. Gogoi, N. Sarma, T. Begum, S. K. Chanda, H. Lekhak, G. N. Sastry, and M. Lal, "Agarwood (Aquilaria malaccensis L.) a quality fragrant and medicinally significant plant based essential oil with pharmacological potentials and genotoxicity," *Ind. Crops Products*, vol. 197, Jul. 2023, Art. no. 116535.

[3] I. D. Thompson, T. Lim, and M. Turjaman, "Expensive, exploited and endangered," in *A Review of the Agarwood-Producing Genera Aquilaria and Gyrinops: CITES Considerations, Trade Patterns, Conservation, and Management*, 2022.

[4] L. Abdulah, R. Susanti, J. S. Rahajoe, T. D. Atikah, S. Subarudi, R. Dewi, I. Heriansyah, M. A. Qirom, K. Rahmawati, and A. Hidayat, "Feasibility of agarwood cultivation in indonesia: Dynamic system modeling approach," *Forests*, vol. 13, no. 11, p. 1869, 2022.

[5] W. Sun, S. Wu, P. Zhu, S. Deng, X. Zhou, X. Liu, J. Zhao, W. Zhang, X. Gao, and X. Chen, "Rapid discrimination of aroma components in natural and artificial agarwood incense by static headspace sampling—Gas chromatography—Mass spectrometry (SHS-GC-MS) with principal component analysis (PCA) and partial least squares—Discrimination analysis (PLS-DA)," *Anal. Lett.*, vol. 55, no. 18, pp. 2829–2845, Dec. 2022.

[6] M. Mohamad, M. S. Najib, S. N. Tajuddin, S. M. Daud, N. F. H. Majid, S. Zaib, and M. F. Zahari, "KNN: Classification of agarwood types in oil and wooden using E-nose," in *Proc. 6th Int. Conf. Elect., Control Comput. Eng.*, 2022, pp. 575-586.

[7] A. Hazarika, J. R. Deka, P. C. Nath, G. W. Sileshi, A. J. Nath, K. Giri, and A. K. Das, "Modelling habitat suitability of the critically endangered agarwood (*Aquilaria malaccensis*) in the Indian east Himalayan region," *Biodiversity Conservation*, vol. 32, no. 14, pp. 4787–4803, Dec. 2023.

[8] K. H. Jung and W. Cho, "Agarwood's domestic and international research trends and literature review of herbal medicine," *Korea J. Herbol.*, vol. 38, no. 6, pp. 1–27, 2023.

[9] B. K. Deep and S. N. Tajuddin, "Agarwood," Tech. Rep.

[10] M. Z. Mustapa, M. A. Alias, S. S. S. A. Azziz, C. F. Wong, M. Ibrahim, R. Yahaya, Y. M. Bakri, and N. A. Rajak, "Agarwood production of aquilaria malaccensis using various inoculants and induction techniques," *Central Asia Caucasus (14046091)*, vol. 23, no. 1, 2022.

[11] A. F. M. Amidon, N. Z. Mahabob, N. Ismail, Z. M. Yusoff, and M. N. Taib, "Agarwood oil quality classification using one versus all strategies in multiclass on SVM model," I-JaMCSIIX Universiti Teknologi MARA Cawangan Melaka Kampus Jasin, Merlimau, Melaka, Tech. Rep. Tel: 062645000, 2021, p. 84.

[12] P. S. Naziz, R. Das, and S. Sen, "The scent of stress: Evidence from the unique fragrance of agarwood," *Frontiers Plant Sci.*, vol. 10, p. 840, Jul. 2019.

[13] A. Faizal, R. R. Esyanti, N. P. Utami, A. W. P. Azar, D. Hermawaty, T. Setiyobudi, A. F. Martin, B. W. Hapsari, and M. Turjaman, "Elicitation of secondary metabolites in *Aquilaria malaccensis* Lam. Callus culture by crude mycelial extract of fusarium solani and Methyl Jasmonate," *Forests*, vol. 14, no. 1, p. 48, Dec. 2022.

[14] M. R. Islam, C. Chakraborty, and S. Banu, "Isolation and characterization of bacteria and fungi associated with agarwood fermentation," *Current Microbiol.*, vol. 79, no. 10, p. 313, Oct. 2022.

[15] P. Shivanand, N. F. Arbie, S. Krishnamoorthy, and N. Ahmad, "Agarwood—The fragrant molecules of a wounded tree," *Molecules*, vol. 27, no. 11, p. 3386, May 2022.

[16] Y. Kang, P. Liu, F. Lv, Y. Zhang, Y. Yang, and J. Wei, "Genetic relationship and source species identification of 58 qi-nan germplasms of aquilaria species in China that easily form agarwood," *PLoS ONE*, vol. 17, no. 6, Jun. 2022, Art. no. e0270167.

[17] J. Dutta, D. Lahon, P. K. Bora, M. Bhuyan, and S. Haldar, "A high performance thin layer chromatography (HPTLC) method for the quality assessment of agarwood (*Aquilaria malaccensis*) oil from northeast India," *Natural Product Res.*, pp. 1–5, Dec. 2022.

[18] K. G. Boyle, G. Napoleone, A. H. Ramsook, R. A. Mitchell, and J. A. Guenette, "Effects of the elevation training mask 2.0 on dyspnea and respiratory muscle mechanics, electromyography, and fatigue during exhaustive cycling in healthy humans," *J. Sci. Med. Sport*, vol. 25, no. 2, pp. 167–172, Feb. 2022.

[19] L. Peeraphong, "Medicinal uses of agarwood," *Chin. Med. Culture*, vol. 4, no. 4, pp. 260–265, 2021.

[20] P. C. Nath, A. J. Nath, G. W. Sileshi, and A. K. Das, "Growth and coppicing ability of the critically endangered agarwood (*Aquilaria malaccensis* Lam.) tree in monoculture and polyculture in north east India," *J. Sustain. Forestry*, vol. 42, no. 9, pp. 947–959, 2022.

[21] S. R. Adhikari, K. Pokhrel, and S. D. Baral, "Economic value of agarwood and its prospects of cultivation," *Int. J. Appl. Sci. Biotechnol.*, vol. 9, no. 1, pp. 23–31, Mar. 2021.

[22] C. S. Tan, N. M. Isa, I. Ismail, and Z. Zainal, "Agarwood induction: Current developments and future perspectives," *Frontiers Plant Sci.*, vol. 10, p. 122, Feb. 2019.

[23] P. Kumar, S. Tripathi, P. K. Rout, S. K. Khare, and S. Naik, "Quality control analysis of high-value agarwood oil by thermogravimetric analysis (TGA)," *Flavour Fragrance J.*, vol. 38, no. 1, pp. 27–36, Jan. 2023.

[24] X. Huang, H. Li, Y. Ruan, Z. Li, H. Yang, G. Xie, Y. Yang, Q. Du, K. Ji, and M. Yang, "An integrated approach utilizing Raman spectroscopy and chemometrics for authentication and detection of adulteration of agarwood essential oils," *Frontiers Chem.*, vol. 10, Dec. 2022, Art. no. 1036082.

[25] A. P. Desa, S. Lee, M. Z. Mustapa, R. Mohamed, and D. Emang, "Trends in the agarwood industry of Peninsular Malaysia," *Malays*, vol. 84, no. 1, pp. 152–168, 2021.

[26] E. E. M. Kahar, B. A. Talip, N. A. M. Fauzi, S. N. Kamarulzaman, F. Zakaria, N. Muhammad, T. N. A. R. Mamat, and H. Basri, "Properties and potential of agarwood hydrosol as a drink: A review," *Food Res.*, vol. 5, no. 3, pp. 29–35, Apr. 2021.

[27] M. A. Haqmi Abas, "Agarwood oil quality classification using support vector classifier and grid search cross validation hyperparameter tuning," *Int. J. Emerg. Trends Eng. Res.*, vol. 8, no. 6, pp. 2551–2556, Jun. 2020.

[28] M. H. Haron, "Agarwood oil quality grading model using selforganizing map (SOM)," Dept. System, Universiti Teknologi MARA, UiTMIR, Shah Alam, Selangor, Tech. Rep., 2020.

[29] N. Safitri, T. Rihayat, S. R. Suryani, R. Saputra, A. Safitri, T. Y. R. Hasnah, K. Y. Putri, and Y. Yunus, "Distillation of agarwood oil (Aquillaria sp) using photovoltaic methods," Tech. Rep., 2020.

[30] A. F. M. Amidon, N. Z. Mahabob, S. M. H. M. Huzir, Z. M. Yusoff, N. Ismail, and M. N. Taib, "The grading of agarwood oil quality based on multiclass support vector machine (MSVM) model," in *Proc. IEEE 18th Int. Colloq. Signal Process. Appl. (CSPA)*, May 2022, pp. 91–95.

[31] N. Z. Mahabob, Z. M. Yusoff, A. F. M. Amidon, N. Ismail, and M. N. Taib, "Modeling of agarwood oil compounds based on linear regression and ANN for oil quality classification," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 11, no. 6, p. 5505, Dec. 2021.

[32] N. Ismail, M. H. F. Rahiman, M. N. Taib, N. A. M. Ali, M. Jamil, and S. N. Tajuddin, "The grading of agarwood oil quality using k-nearest neighbor (k-NN)," in *Proc. IEEE Conf. Syst., Process Control (ICSPC)*, Dec. 2013, pp. 1–5.

[33] M. H. Haron, "Grading of agarwood oil quality based on its chemical compounds using self organizing map (SOM)," *Int. J. Emerg. Trends Eng. Res.*, vol. 8, no. 7, pp. 3728–3736, Jul. 2020.

[34] N. Z. Mahabob, A. F. M. Amidon, N. Ismail, S. M. H. M. Huzir, Z. M. Yusoff, M. N. Taib, S. N. Tajuddin, and N. M. Ali, "A study on ANN performance towards three significant compounds of high quality agarwood oil," in *Proc. IEEE 18th Int. Colloq. Signal Process. Appl. (CSPA)*, May 2022, pp. 116–120.

[35] N. S. Ismail, N. Ismail, M. H. F. Rahiman, M. N. Taib, N. A. M. Ali, and S. N. Tajuddin, "Polynomial tuned kernel parameter in SVM of agarwood oil for quality classification," in *Proc. IEEE Int. Conf. Autom. Control Intell. Syst. (I2CACIS)*, Oct. 2018, pp. 77–82.

[36] N. Ismail, S. L. I. Samawi, Z. M. Yusoff, M. H. F. Rahiman, S. N. Tajuddin, and M. N. Taib, "Decision tree on branch node observations for significant compounds of agarwood oil different grades," *J. Tomogr. Syst. Sensor Appl.*, vol. 6, no. 1, pp. 43–51, 2023.

[37] S. O. Sada, "Improving the predictive accuracy of artificial neural network (ANN) approach in a mild steel turning operation," *Int. J. Adv. Manuf. Technol.*, vol. 112, nos. 9–10, pp. 2389–2398, Feb. 2021.

[38] Z. Y. Shnain, A. A. Alwaiti, M. K. Rashed, and Z. M. Shakor, "Experimental and data-driven approach of investigating the effect of parameters on the fluid flow characteristic of nanosilica enhanced two phase flow in pipeline," *Alexandria Eng. J.*, vol. 61, no. 2, pp. 1159–1170, Feb. 2022.

[39] J. C. Tellez Gaytan, K. Ateeq, A. Rafiuddin, H. M. Alzoubi, T. M. Ghazal, T. A. Ahanger, S. Chaudhary, and G. K. Viju, "AI-based prediction of capital structure: Performance comparison of ANN SVM and LR models," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–13, Sep. 2022.

[40] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, A. M. Umar, O. U. Linus, H. Arshad, A. A. Kazaure, U. Gana, and M. U. Kiru, "Comprehensive review of artificial neural network applications to pattern recognition," *IEEE Access*, vol. 7, pp. 158820–158846, 2019.

[41] H. Shareef, M. S. Ahmed, A. Mohamed, and E. Al Hassan, "Review on home energy management system considering demand responses, smart technologies, and intelligent controllers," *IEEE Access*, vol. 6, pp. 24498–24509, 2018.

[42] M. A. Rahu, A. F. Chandio, K. Aurangzeb, S. Karim, M. Alhussein, and M. S. Anwar, "Toward design of Internet of Things and machine learning-enabled frameworks for analysis and prediction of water quality," *IEEE Access*, vol. 11, pp. 101055–101086, 2023.

[43] I. Krop, Y. Takahashi, T. Sasaoka, H. Shimada, A. Hamanaka, and J. Onyango, "Assessment of selected machine learning models for intelligent classification of flyrock hazard in an open pit mine," *IEEE Access*, vol. 12, pp. 8585–8608, 2024.

[44] J. K. Janga, K. R. Reddy, and K. V. N. S. Raviteja, "Integrating artificial intelligence, machine learning, and deep learning approaches into remediation of contaminated sites: A review," *Chemosphere*, vol. 345, Dec. 2023, Art. no. 140476.

[45] N. Z. Mahabob, *Intelligent Grading of Agarwood Essential Oil Quality Using Artificial Neural Network (ANN)*. Shah Alam, Malaysia: Universiti Teknologi MARA (UiTM), 2022.

[46] Y. Ren and D. Goldfarb, "Efficient subsampled Gauss–Newton and natural gradient methods for training neural networks," 2019, *arXiv:1906.02353*.

[47] K. Huang and X. Fu, "Low-complexity proximal Gauss–Newton algorithm for nonnegative matrix factorization," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2019, pp. 1–5.

[48] X. Li, T. Bi, Z. Wang, L. Xu, and Y. He, "Damped Gauss–Newton based online ranging for point extraction from low SNR and high overlapping waveforms," *Measurement*, vol. 199, Aug. 2022, Art. no. 111479.

[49] A. C. Cinar, "Training feed-forward multi-layer perceptron artificial neural networks with a tree-seed algorithm," *Arabian J. Sci. Eng.*, vol. 45, no. 12, pp. 10915–10938, Dec. 2020.

[50] M. A. El-Brawany, D. A. Ibrahim, H. K. Elminir, H. M. Elattar, and E. A. Ramadan, "Artificial intelligence-based data-driven prognostics in industry: A survey," *Comput. Ind. Eng.*, vol. 184, Oct. 2023, Art. no. 109605.

[51] E. Boido, L. Fariña, F. Carrau, D. Cozzolino, and E. Dellacassa, "Application of near-infrared spectroscopy/artificial neural network to quantify glycosylated norisoprenoids in tannat grapes," *Food Chem.*, vol. 387, Sep. 2022, Art. no. 132927.

[52] H. Özdoğan, Y. Ali Üncü, O. Karaman, M. Sekerci, and A. Kaplan, "Estimations of giant dipole resonance parameters using artificial neural network," *Appl. Radiat. Isot.*, vol. 169, Mar. 2021, Art. no. 109581.

[53] F. Sherwani, B. S. K. K. Ibrahim, and M. M. Asad, "Hybridized classification algorithms for data classification applications: A review," *Egyptian Informat. J.*, vol. 22, no. 2, pp. 185–192, Jul. 2021.

[54] A. Foroutan, S. Basumallik, and A. Srivastava, "Estimating and calibrating DER model parameters using Levenberg–Marquardt algorithm in renewable rich power grid," *Energies*, vol. 16, no. 8, p. 3512, Apr. 2023.

[55] L. Wu, D. He, B. Ai, J. Wang, H. Qi, K. Guan, and Z. Zhong, "Artificial neural network based path loss prediction for wireless communication network," *IEEE Access*, vol. 8, pp. 199523–199538, 2020.

[56] V. Amaratunga, L. Wickramasinghe, A. Perera, J. Jayasinghe, and U. Rathnayake, "Artificial neural network to estimate the paddy yield prediction using climatic data," *Math. Problems Eng.*, vol. 2020, pp. 1–11, Jul. 2020.

[57] N. Z. Mahabob, Z. M. Yusoff, A. F. M. Amidon, N. Ismail, and M. N. Taib, "A novel application of artificial neural network for classifying agarwood essential oil quality," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 6, p. 6645, Dec. 2022.

SITI MARIATUL HAZWA MOHD HUZIR was born in Malaysia, in May 1997. She received the Bachelor of Engineering degree (Hons.) in electronics engineering from Universiti Teknologi MARA (UiTM), Shah Alam, Malaysia, in 2021. She is currently pursuing the Ph.D. degree in electrical engineering with UiTM, Johor Branch, Pasir Gudang Campus. Her current research interests include artificial intelligence, automation, robotics, modeling, and machine learning.

ANIS HAZIRAH 'IZZATI HASNU AL-HADI was born in Malaysia, in July 1998. She received the B.Eng. degree (Hons.) in electronics engineering from Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia, in 2022, where she is currently pursuing the master's degree with the School of Electrical Engineering. Her research interests include artificial intelligence and machine learning.

ZAKIAH MOHD YUSOFF (Member, IEEE) received the B.Eng. and Ph.D. degrees in electrical engineering from UiTM Shah Alam, in 2009 and 2014, respectively. In May 2014, she joined as a Teaching Staff with UiTM Pasir Gudang, where she is currently a Senior Lecturer. Her current research interests include process control, system identification, and essential oil extraction systems.

NURLAILA ISMAIL (Member, IEEE) received the Ph.D. degree in electrical engineering from Universiti Teknologi MARA, Malaysia. She is currently a Senior Lecturer with the School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Malaysia. Her research interests include advanced signal processing and artificial intelligence.

MOHD NASIR TAIB (Senior Member, IEEE) received the Ph.D. degree from UMIST, U.K. He is currently a Senior Professor with Universiti Teknologi MARA (UiTM). He also heads the Advanced Signal Processing Research Group, School of Electrical Engineering, College of Engineering, UiTM. He has been a very active researcher and over the years has authored and/or coauthored many papers published in refereed journals and conferences. His research interests include control and instrumentation.