IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 10, 2021, accepted January 19, 2021, date of publication January 26, 2021, date of current version February 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3054816

Examining the Forecasting Movement of Palm Oil Price Using RBFNN-2SATRA Metaheuristic Algorithms for Logic Mining

SHEHAB ABDULHABIB SAEED ALZAEEMI AND SARATHA SATHASIVAM
School of Mathematical Sciences, Universiti Sains Malaysia (USM), Penang 11800, Malaysia

Corresponding author: Saratha Sathasivam (saratha@usm.my)

This work was supported in part by the Fundamental Research Grant Scheme (FRGS) (203/PMATHS/6711689) by the Ministry of Higher Education Malaysia, and in part by Universiti Sains Malaysia.

ABSTRACT RBFNN with different algorithms and the logic mining method for forecasting constitute the most significant tools and techniques, which are used to demonstrate the economic growth in the country. Upon using monthly data spanning from Jan 2016 to March 2020 for the manufacturing of palm oil, the results mainly revealed that RBFNN-2SATRAAIS is the most accurate and efficient model compared to RBFNN-2SATRAPSO and RBFNN-2SATRAGA in forecasting the price of palm oil. RBFNN-2SATRAAIS had the highest average overall accuracy (90.476190%), followed by RBFNN-2SATRAPSO (85.71%), and RBFNN-2SATRAGA (76.19%). The results also showed that the spot price of palm oil is highly influenced by the total exports and imports, as well as the production of palm oil, its end-stocks, and Malaysia's real, effective exchange rate. By using the data mining technique based on the energy minimization technique, the logical mining task was carried out. The empirical findings provided useful insights into decision-making and policy implementations, including the formulation of strategies to help the industry in dealing with constant price fluctuations and, thereby, enabling the Malaysian palm oil industry to continue its domination over the international market.

INDEX TERMS Palm oil prices, economic growth, forecasting, artificial immune system, particle swarm optimization, genetic algorithm, radial basis functions neural network, 2 satisfiability, 2 satisfiability reverse. Analysis; logic mining.

I. INTRODUCTION

There are many industries, which play a pivotal role in the Malaysian economy. Among these industries is the Malaysian palm oil industry that enhances the country's economic-social level. This has been demonstrated by the contribution of the significant palm oil industry in Malaysia. For illustration, as a key contributor in 2015, palm oil represents the fourth main contributing product to the country's national income, accounting for 63 billion RM of Malaysia's gross national income [1]. Malaysia has had substantial expertise in the palm oil industry for more than 100 years. Malaysia has a comparative success in the international market and, thus, it has become a market leader in terms of productivity research and development. The government of Malaysia has superintended and carried out several frameworks and initiatives to intensify the predominance of the essential Malaysian industry of palm oil in the international market, thereby obliquely enhancing the country's income. Besides being a key income source for Malaysia, Malaysia's Palm Oil Council announced that the Malaysian palm oil industry contributes to the influential employment opportunities for more than 650,000 labor and smallholders and more faraway 3 million people in Malaysia, whose sustenance is very much dependent on this industry [2]. However, there is fluctuation in prices of palm oil between down and rise during the previous period as shown in FIGURE 1. This scenario frequently happens along the period and the gap of the price falls is quite big. Such a fluctuation trend in the prices of palm oil has increased concerns for those who are dealing with uncertainties and risks in the oil palm trade and has influenced the smallholder's income, which will further affect the country's revenue in the long run.

The associate editor coordinating the review of this manuscript and approving it for publication was Chien-Ming Chen.

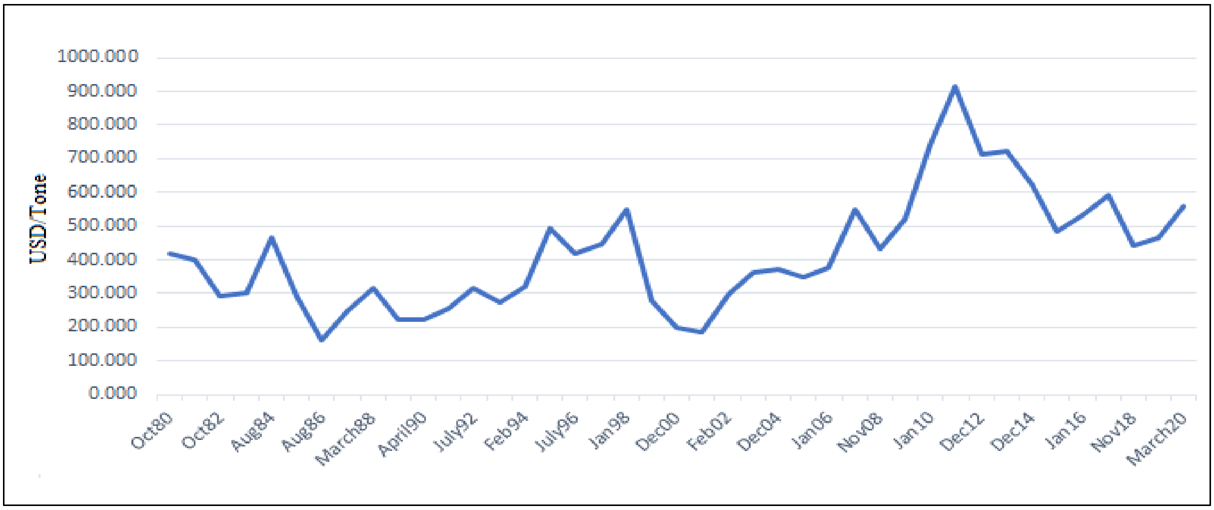

FIGURE 1. **The trend of Malaysia's palm oil price per year.**

Therefore, it is crucial to understand the palm oil price behavior and its affecting attributes in order to ensure a stable future price and to help in making wise decisions. Therefore, it is imperative to examine the determinants of attributes of the price change to properly anticipate these fluctuations in the price of palm oil over time. The techniques of forecasting were examined in previous studies, but researchers have implemented conventional forecasting techniques [3]–[5] and these techniques did not achieve more consistent findings. New techniques of forecasting exist nowadays, in which, more reliable findings can be yielded such as using Artificial Neural Networks (ANN) models in forecasting [6]. A recent study by [7] showed the ability of the recurrent neural network (RNN) as the building block for the model of ability-aware Person Job Fit Neural Network (APJFNN) in training an industrial data set in China. The proposed model has recorded better accuracy compared with the state-of-the-art approaches like decision tree, linear regression, and gradient boosting decision tree. Since most of the aforementioned logic mining techniques integrate the statistical measures, an alternative method will be appropriate to facilitate the learning and testing phase of the recruitment data set.

This paper aims to consider logic mining in the palm oil price behavior so that a novel method is proposed to identify the relationship among the attributes of the monthly price of palm oil and how it affects the output of the palm oil industry in Malaysia. This can be achieved by using the Radial Basis Function Neural Network and the logic mining method.

Radial Basis Function Neural Network (RBFNN) has wide-ranging uses for countless types of problems in business, industry, and science [8]. One major use of RBFNN involves time series forecasting [9]. Many successful applications suggested that RBFNN can be a promising alternative tool for both forecasting researchers and practitioners [10]. Research on artificial neural networks (ANN), specifically the radial basis function neural network is still primitive. Various metaheuristics have been used to enhance the performance of the neural network various ANN models for different purposes. According to the No Free Lunch theorem (NFL), there has been no metaheuristics algorithm, which can perform better than all other algorithms in all optimization problems [11]. Motivated by this reason, different algorithms were proposed to train RBFNN-2SATRA to compare the efficiency and effectiveness of these algorithms and obtain the best algorithm to train RBFNN-2SATRA. The advantages of metaheuristics algorithms involve flexibility, self-adaptation, conceptual simplicity, and the ability to search for a global optimum rather than a local one [12]–[14]. This paper aims to explore the training capacity in RFBNN-2SATRA. In this regard, research is underway for other potential metaheuristics, and they will be explored in the future.

The introduced logic mining method has the capability of extracting the logical rule between neurons. In [15], a logic mining method in ANN has been presented by applying the 3 Satisfiability Reverse Analysis method. To improve the method, [16] proposed a logic mining method by implementing in RBFNN and 2 Satisfiability. The proposed model showed its capability of extracting the logical rule between neurons. The power of logic mining can be updated when the new data is introduced to the database. The newly introduced data will, therefore, be updated in the data sets, which will contribute to the new relationship or trend of data. Thus, it demonstrates the capability of our logic mining models to work with the Malaysian palm oil price behavior data with different dimensions and instances. Logic mining techniques are widely used in several research domains and provide useful results to guide the decision-makers [11], [17]. The results of this technique, besides being of interest, can guide decision-makers. It is worth mentioning that the induced logical rule of logic mining is always converged to global minimum energy [18]. The successful development of effective and efficient logic mining by applying algorithms can provide enormous benefits to an organization from a business standpoint. These benefits include reduced costs due to more accurate control, more accurate future predictions, more effective fault detection and prediction, fraud detection and control, and automation of repetitive human tasks. According to Sikora and Piramuthu [19], they presented the design of a more effective and efficient GA to the logic mining techniques that use the concepts of self-adaptive feature selection together with a wrapper feature selection method based on Hausdorff distance measure. In another development to logic mining, Yuan [20] presented an effective algorithm called PSO with the logic mining for Mass Sensor Networks. Evolutionary algorithms for improving logic mining have become a very popular research topic in recent years. Among many proposed algorithms, the three algorithms that are very similar and popular include GA, PSO, and AIS. While GA is more well-established because of its much earlier introduction, the more recent PSO and AIS algorithms have started to attract more attention. Many papers were published based on these algorithms. However, most assessments have been made empirically with specific cases on a particular application domain, and the two key performance indicators that were commonly used for comparison are the solution quality and solution time. This paper aims to conduct a general qualitative comparison of the three evolutionary algorithms based on logic mining. By implementing metaheuristics such as GA, PSO, and AIS, the proposed logic mining can, therefore, reduce the error that leads to suboptimal induced logic [12].

The implementation of the 2SAT in the Reverse Analysis method in extracting the data set is still limited. In fact, there are no significant attempts and recent developments in extracting attributes as the factors in terms of systematic logical representation that contributed to the Malaysian palm oil price behavior data. This is the main motivation behind introducing the systematic variant of logical rule, 2SATRA as the logical representation to properly represent the extracted logic for the Malaysian palm oil price behavior data cases. The previous work on the reverse analysis method by incorporating Horn logical rule in the neural network was introduced by [21] in processing the monitoring stock in a departmental store performance dataset. However, the behavior of the data is represented by using a non-systematic logical rule, where the redundancy of attributes was considered. In the case of the Malaysian palm oil price behavior data, the attributes and instances consistently vary every month. Therefore, the systematic logical representation such as 2SATRA is crucial to analyze the pattern of the Malaysian palm oil price behavior data set. To sum up, the primary aim of the present study involves developing a logical rule to explore the relationship among the candidate's attributes, which contributes to making proper anticipation regarding the fluctuations in the Malaysia palm oil price over time. To address the problem, the method of Reverse Analysis referred to as the 2 Satisfiability-based Reverse Analysis method (2SATRA) is utilized for the data extraction method for the Malaysian palm oil price behavior data set. The proposed method involves less complicated mathematical formulas and theorems. In order to use a Radial Basis Function Neural Network (RBFNN) with 2 Satisfiability based Reverse Analysis method (2SATRA) to make predictions for a realistic problem, the network must be trained. RBFNN-2SATRA has been trained by using three algorithms, namely Artificial Immune System algorithm (AIS), Particle Swarm Optimization algorithm (PSO), in addition to Genetic Algorithm (GA) to extract the best logic for the Malaysian palm oil price behavior data set. In this paper, three technologies were proposed, namely Radial Basis Function Neural Network-based 2 Satisfiability Reverse Analysis with Artificial Immune System algorithm (RBFNN-2SATRAAIS), Radial Basis Function Neural Network-based 2 Satisfiability Reverse Analysis with Genetic Algorithm (RBFNN-2SATRAGA), and Radial Basis Function Neural Network-based 2 Satisfiability Reverse Analysis with Particle Swarm Optimization algorithm (RBFNN-2SATRAPSO) for extracting the best logic mining for the Malaysian palm oil price.

This work advances the following contributions: (1) To convert the Malaysian Palm Oil Price data set into a systematic form based on 2 Satisfiability (2SAT) logic. (2) To apply the 2 Satisfiability Reverse Analysis (2SATRA) method as an alternative approach in extracting the relationships between the factors or attributes, which contribute to forecasting the crude palm oil price in the future in Malaysia. (3) To assess the capability and accuracy of the 2SATRA approach based on RBFNN. (4) To utilize various algorithms like GA, PSO, and AIS to identify the relationships between the factors or attributes, which contribute to forecasting the crude palm oil price in the future in Malaysia.

The performance evaluation metrics were adopted to evaluate the effectiveness of three algorithms used to train the RBFNN-2SATRA to extract the best logic mining for the Malaysian palm oil price data set with a varied Hidden Neurons' number in the RBFNN's hidden layer. Thus, to forecast the future price of crude palm oil, all algorithms were utilized. The results from each algorithm were then compared to identify the best model in forecasting the crude palm oil price. This study is significant research since Malaysia is one of the major exporting countries of palm oil and, therefore, the results may provide useful insights into ensuring the competitiveness of Malaysia's palm oil sector in the global market. FIGURE 1 shows the trend of Malaysia's Palm Oil Price per year.

II. TWO SATISFIABILITY REPRESENTATION (2SAT)

Two Satisfiability (2SAT) can be defined as a specific logical representation. It strictly comprises 2 variables for each clause [22]. The 2 SAT properties are summarized as in the following:

i. A specific set of m logical variables, $x_1, x_2, \ldots, x_m$. Each of the variables can store a given binary value $x_i \in \{1, 0\}$, exemplifying correspondingly TRUE/FALSE.

ii. In x_i every variable is a given set of literals, whereby positive and negative literals are x_m and $\neg x_m$, correspondingly.

iii. Consisting of a specific set of n distinct clauses, $C_1, C_2, C_3, \ldots, C_n$. Each of the C_i is linked by logical AND ($\wedge$), and every 2 literals form a given single C_i linked by a logical OR ($\vee$).

Upon using property (i) throughout (iii), 2SAT formulation's definition is explicitly defined or P_{2SAT}:

$$P_{2SAT} = \bigwedge_{i=1}^{n} C_i = \bigvee_{j=1}^{n} \left(x_{ij}, y_{ij}\right) \tag{1}$$

An example of P_{2SAT} is:

$$P_{2SAT} = (A \vee B) \wedge (C \vee D) \wedge (\neg E \vee F) \tag{2}$$

The formulation of both P_{2SAT} is characterized in the Conjunctive Normal Form (2CNF). This is due to the nature of CNF Satisfiability, which is conserved compared to forms like the Disjunctive Normal Form (DNF).

This paper has provided information concerning the price datasets of the Malaysian palm oil in terms of attributes, i.e., variables in the P_{2SAT}, and they are the ANN symbolic rule.

III. RADIAL BASIS FUNCTION NEURAL NETWORK

Radial Basis Function Neural Network (RBFNN) can be defined as a specifically modified feed-forward neural network of a given hidden, as well as an interconnected layer, which has been established by [23] and [24]. As opposed to other networks, RBFNN enjoys further integrated architecture and structure. Regarding its structure, RBFNN comprises three layers of neurons for specific computation purposes. The m neurons in the input layer are input data transferred to a system. The parameters (i.e., center and width), during the training phase, are computed in the given hidden layer. The obtained parameters are utilized for calculating the given output weight in the given output layer. In order for dimensionality to be reduced from an input to an output layer, Gaussian activation's function is provided, whereby the Hidden Neuron's Gaussian activation's function $\varphi_i(x)$ in RBFNN involves [25]:

$$Q(x) = \frac{\left\| \sum_{j=1}^{N} w'_{ji} x_j - c_j \right\|^2}{2\sigma_j^2} \tag{3}$$

$$\varphi_i(x) = e^{-Q(x)}, \tag{4}$$

whereby c_j, σ_j refer to center, as well as the width of the given Hidden Neuron, correspondingly. Based on the present case, x_j refers to the given input value of N the input neurons. Consequently, Euclidean norm $\|\|$ from the neuron i throughout the neuron j involves:

$$\left\| \sum_{j=1}^{N} w'_{ji} x_j - c_i \right\| = \sqrt{\sum_{m=1}^{N} \left(\sum_{j=1}^{N} w'_{ji} x_j - c_i \right)^2} \tag{5}$$

whereby w'_{ji} refers to the given input weight amid input neuron j, as well as Hidden Neuron i. Structurally, x_j refers to the given input data in a training set, as well as the given Hidden Neuron i. c_i refers to the Hidden Neuron's center. The following equation provides the RBFNN's final output $f(w_i)$:

$$f(w_i) = \sum_{i=1}^{N} w_i \varphi_i(x_k), \tag{6}$$

whereby $f(w_i) = (f(w_1), f(w_2), \ldots, f(w_N))$, referring to the RBFNN output value and the given output weight provided by $w_i = (w_1, w_2, \ldots, w_N)$.

The goal of RBFNN involves obtaining optimum weights w_i, thereby satisfying the chosen value of output. A Hidden Neuron in RBFNN provides a group of functions that represent a given input pattern spanned by a Hidden Neuron [26].

IV. 2 SATISFIABILITY REVERSE ANALYSIS METHOD (2SATRA) IN RBFNN

The main impetus of logic mining involves extracting valuable logical rule from the given data sets. This section formulates an improved Reverse Analysis method, which is the 2 Satisfiability Reverse Analysis method of 2SATRA. 2SATRA refers to a specific logic mining method; it uses the RBFNN-2SAT models for extracting logical rules from the given data set. The application of 2 Satisfiability enhanced Reverse Analysis approach in hybrid RBFNN networks can be shown using this algorithm:

Step 1: Assuming the binary learning, as well as the testing Malaysian Palm Oil Price dataset with the outcome of P_{learn} (60%) and P_{test} (40%) converting all the raw Malaysian Palm Oil Price dataset to binary [15], [18].

Step 2: Initializing specific input data, the width, as well as the center of the neurons; assigning all binary data neurons depending on the first conducted step.

Step 3: Segregating a two neurons' set for each clause $L_1, L_2, \ldots, L_n$ leading to $P_{learn} = 1$.

Step 4: Obtaining P_{best} by relating the 2SAT clauses frequency in the complete learning Malaysian Palm Oil Price dataset.

Step 5: Checking clause output weight of hidden layer of P_{best} by using RBFNN-2SATRAGA, RBFNN-2SATRAPSO, and RBFNN-2SATRAAIS.

Step 6: Extricating the best output weight W_i of P_{best}.

Step 7: Finding the neurons' final state by calculating the equivalent output of the RBFNN-2SAT following [27] in the following equation:

$$\operatorname{sgn}(f(w_i)) = \begin{cases} 1, & f(w_i) \leq 0 \\ 0, & Otherwise, \end{cases} \tag{7}$$

where $f(w_i)$ refers to an output value, as well as an output weight of RBFNN, i.e., w_i.

$$f(w_i) = \sum \varphi_i(x) w_i, \tag{8}$$

where φ_i refers to the activation function of input x_i in a hidden layer, as well as W_i, which refers to the weight amid

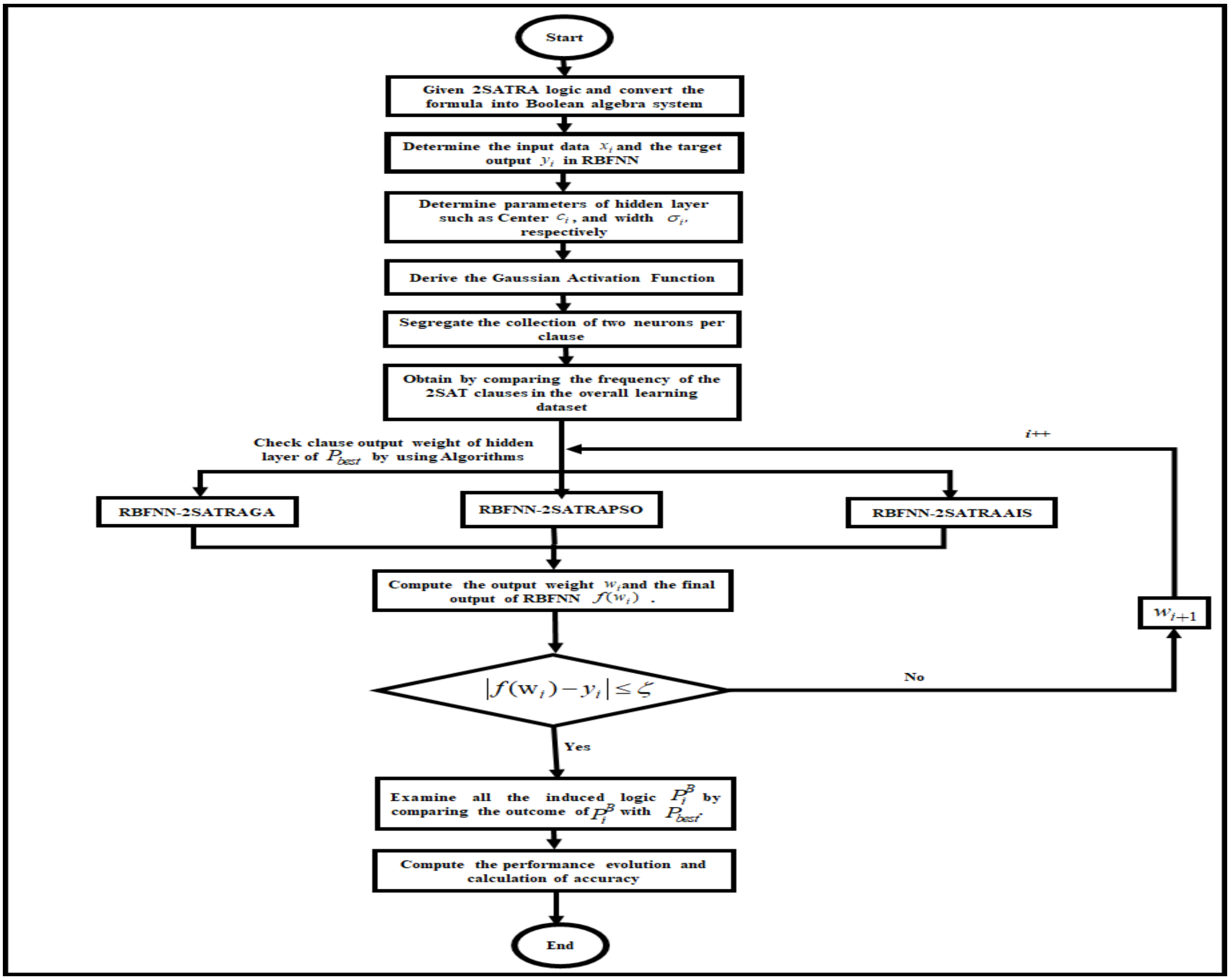

FIGURE 2. Methodology flowchart.

input data in a hidden layer, as well as output data in an output layer.

Step 8: Inducing the entire probable 2SAT logic $P_1^B, P_2^B, \ldots, P_n^B$ from neurons' states.

Step 9: Examining the entire induced logic P_i^B by contrasting the given outcome for P_i^B with the P_{test}.

Step 10: Obtaining the entire performance evolution, as well as calculating accuracy.

2SATRA can be defined as a method, which uses the advantageous feature of the RBFNN, as well as the 2 Satisfiability logic or RBFNN-2SATRA. 2SATRA represents a feasible method that helps to extract the most ideal logical rule, which governs the performance of the data set. The complete flowchart of FIGURE 2 shows the summarized methodology of this study.

V. 2SATRA IN PALM OIL DATA SET

Palm Oil (PO) data is a data set, which is obtained from the Malaysian palm oil industry in Malaysia. In this regard, [22] studied a few variables only in their study without considering other important variables, which may exert a considerable impact on the prices of palm oil like Crude Oil (i.e., petroleum), as well as End-Stocks of Palm Oil, in addition to the Palm Oil Total Exports, the Production of Palm Oil, with the Palm Oil Total Imports. This paper aims to bridge the gap by including relevant determinants such as Total Exports (MT) of Palm Oil, Total Imports (MT) of Palm Oil, Palm Oil production, Palm Oil End-Stocks, Malaysia's Real Effective Exchange Rate, as well as Crude Oil (petroleum) - the Monthly Price- USA in order to understand the Malaysian monthly price for palm oil (USA/Ton) movement. In contrast to most previous literature that used a univariate approach in forecasting the palm oil price, this study uses a multivariate approach, in which, several new explanatory variables are included in forecasting the movement of the Malaysian monthly price of palm oil (USA/Ton). The logic mining approach via 2SATRA provides a solid logical rule as a depiction of the Malaysian monthly price of the palm oil data set to enhance the export performance of the Malaysian palm oil. This has been utilized to generate a logical rule

TABLE 1. Data description.

Variables	Description
P_{2SAT}	The monthly price for palm oil (USA/Ton)
A	Total Exports (MT) of Palm Oil
B	Total Imports (MT) of Palm Oil
C	Production of Palm Oil
D	End-Stocks of Palm Oil
E	Malaysia Real Effective Exchange Rate
F	Crude Oil (petroleum)-Monthly Price-USA

for demonstrating the data's behavior. The data set are as illustrated in TABLE 1.

whereby MT means Metric Ton. In terms of methodology, this study uses three different algorithms to enhance the export performance of Malaysian palm oil and to forecast the spot price in the future. The algorithms include the Artificial Immune System algorithm (AIS), the Particle Swarm Optimization algorithm (PSO), and the Genetic Algorithm (GA). Each of these algorithms' details is discussed as follows:

A. RADIAL BASIS FUNCTION NEURAL NETWORK-BASED 2 SATISFIABILITY REVERSE ANALYSIS WITH GENETIC ALGORITHM (RBFNN-2SATRAGA)

The Genetic Algorithm refers to a typical metaheuristic algorithm, which is utilized to tackle varied optimization problems. Assuming a specific finite solution space, GA structure is classified into local search, as well as global search [28]. The strings' populations in GA, i.e., chromosomes are characterized in relation to the solutions to an optimization problem [29]. The chromosome's quality is represented by a fitness value, and at each generation, each chromosome's fitness value is assessed. The best fitness will be chosen as a given final solution. Chromosomes improve their fitness by applying three operators, including crossover and selection, in addition to mutation. Thus, the crossover can promote information exchange among chromosomes. In this regard, Hamadneh *et al.* [26] utilized GA for identifying the Hidden Neurons' centers, width, along with the Hidden Neuron's number via lessening the amount of the total error of the actual outputs, in addition to favorite outputs. Throughout the selection, therefore, several chromosomes will be chosen from the existing population, which depends upon the fitness value. Also, the mutation is added so that the chromosomes' genetic diversity is created. GA has been utilized in this work for optimizing the RBFNN-2SATRA output weight via minimizing the error of training. The GA implementation in RBFNN is referred to as RBFNN-2SATRAGA. In the RBFNN-2SATRAGA, the output weight is calculated by GA. The following steps are involved in the RBFNN-2SATRAGA.

Step 1: Initialization of Population

Output weights, which are characterized by a given chromosome, are initialized. The following are the chromosomes' representations:

$$w_i = (w_1, w_2, w_3, \ldots, w_N) \tag{9}$$

A population possesses N_{pop} chromosomes, comprising N_N randomly output weights, and the aim involves minimizing objective function:

$$f_{GA}(w_i) = \sum_{i=1}^{j} w_i \varphi_i(x) \tag{10}$$

whereby $f_{GA}(w_i)$ refers to an objective function in the model of RBFNN-2SATRAGA.

Step 2: Computation of Fitness

The fitness of each of the chromosomes can be computed through the RBFNN-2SATRA basis function, which is utilized in the present work. It can be displayed in this equation as follows:

$$fit_i = \frac{1}{(1 + f_{GA}(w_i))}, \quad 0 \le fit \le 1 \tag{11}$$

whereby $f_{GA}(w_i)$ refers to an objective function; fit_i which refers to fitness value.

Step 3: Selection

The chromosomes' arrangement is in a specific descending order according to the fitness function value, whereby the optimum chromosomes (i.e., of the highest value of fitness) will be kept only, whereas the remaining ones will be rejected. The selection probability, p_i for each of the chromosomes, can be computed by utilizing this following equation:

$$p_i = \frac{fit_i}{\sum_{i=0}^{n} fit_i} \tag{12}$$

whereby fit_i refers to the given fitness value.

Step 4: Crossover

All through the phase of crossover, information from a given parent is randomly exchanged to produce an offspring of diverse genetic composition, and the crossover's location is randomly chosen. The phase of crossover determines the cross-population number based on the rate of crossover. Assume a couple of parents w_k, as well as w_m, the offspring w_i^{new} generated by these equations [12], [30]:

$$w_i^{new} = \begin{cases} w_m + r(w_k - w_m), & p_i, \ i = 1, 2, 3, \ldots n \\ w_m, & 1 - p_i \end{cases} \tag{13}$$

whereby p_i refers to probability, r refers to the rate of crossover, w_k refers to the higher probability chromosome, w_m refers to the lower probability chromosome, whereas parameter k has been selected by this equation:

$$k = \begin{cases} rand(m, n), & p_i \\ m, & 1 - p_i \end{cases} \tag{14}$$

whereby $k + m = n$ and $k > m$. Value k is disseminated amid k, as well as m uniformly.

Step 5: The Mutation Phase

Throughout this phase, the chromosome's information is arbitrarily assigned in the pre-identified range (i.e., often identified by users). The mutation can be predicted to generate a new chromosome breed based on this equation:

$$w_m^{new} = \begin{cases} rand\,(-5, 5)\,, & rand(0, 1) < \tau \\ w_i, & rand(0, 1) \geq \tau \end{cases} \tag{15}$$

where w_m^{new} refers to a novel chromosome from the phase of mutation when $\tau \in [0,\ 1]$.

Step 6: The Termination Phase

GA can iterate to 10000 Generations. In case a given criterion of a solution's termination is fulfilled, calculation of the algorithm will be terminated, or it will return to the previous Step 2 with $i = i + 1$. The final output of RBFNN-2SATRA will be the chromosome, which has the RBFNN-2SATRAGA's ideal output weights.

B. RADIAL BASIS FUNCTION NEURAL NETWORK-BASED 2 SATISFIABILITY REVERSE ANALYSIS WITH PARTICLE SWARM OPTIMIZATION ALGORITHM (RBFNN-2SATRAPSO)

PSO signifies a given class of the iterative swarm-based searching algorithm. It enjoys a wide-ranging utilization, i.e., it can be utilized as a specific learning algorithm, or it can be used as a universal optimization. The innovative work of PSO was performed by Eberhart and Kennedy [31] via the mathematical modelling of socio-behavioral features belonging to birds' gathering and schooling of fish in each population. The most remarkable feature of PSO involves adjustable free parameters that facilitate implementation and optimization. PSO has particularly used a robust searching process via impending the ideal particle in the given solution space [32]. Potential solutions, i.e., particles are flying over each searching space by following the existing superlative particles. Also, the change in the particles' position occurs in the PSO algorithm, whereby it is necessary to search for the ideal particle. PSO algorithm has been adopted in this study to optimize an output weight amidst Hidden Neurons, in addition to enhance the RBFNN-2SATRA output neurons. These steps are the main steps in the RBFNN-2SATPSO model:

Step 1: Initialization

Initialization of the population particles, where $w_i = (w_1, w_2, w_3, \ldots, w_n)$, and its position $x_i = (x_1, x_2, x_3, \ldots, x_n)$.

Step 2: Evaluate initial populations

Evaluate initial populations using the objective function $f_{PSO}\,(w_i)$ of each particle in the swarm.

$$f_{PSO}\,(w_i) = \sum_{i}^{n} w_i\,\varphi_i\,(x) \tag{16}$$

where $w_i \in \mathbb{R}$ refers to the output weight and $\varphi_i(x)$ is the Gaussian Activation Function in RBFNN-2SATRA.

Step 3: Set to the objective value

Find the best values p_i^{best} from the minimum value of an objective function. Compare p_i^{best} of particles with each other and update the swarm global best location with the greatest objective function (g_i^{best}).

Step 4: Update the particle velocity

The particle updates its velocity by the following equation:

$$w_{i+1} = \Omega\, w_i + \varepsilon_1\, rand_1\, \left(p_i^{best} - x_i\right) + \varepsilon_2\, rand_2\, \left(g_i^{best} - x_i\right) \tag{17}$$

The parameter $\Omega = 0.6$ is called the inertia weight, whereas $\varepsilon_1 = \varepsilon_2 = 2$ are acceleration constants, $rand_1 = rand_2$ are randomly sampled in the range [0, 1], p_i^{best} refers to the personal optimum position, which is accomplished by the particle (output weight) of the first swarm and g_i^{best} refers to the global optimum position, which is accomplished by the particles (output weights) of the second swarm.

Step 5: Termination

PSOA iterates to 10000 iterations. When a specific solution termination criterion is achieved, the algorithm's calculation is terminated, or it will update the position of the new particle by the following equation:

$$x_{i+1} = x_i + w_{i+1} \tag{18}$$

where x_i is the initial position. The final output of RBFNN-2SAT is the optimal output weights of RBFNN-2SATPSOA.

C. RADIAL BASIS FUNCTION NEURAL NETWORK-BASED 2 SATISFIABILITY REVERSE ANALYSIS WITH ARTIFICIAL IMMUNE SYSTEM ALGORITHM (RBFNN-2SATRAAIS)

Having enthused by nature, the non-traditional optimization methods have been popularly used in combinatorial optimization. The algorithm of the artificial immune system is an inspired technique by the immune system. AIS refers to an adaptive system stimulated by theoretical immunology, as well as the immune functions and principals, along with models applied to intricate problem domains [33]. AISs can be used in areas like biological modelling and the computer network's security, detection of viruses, robotics, data mining, scheduling, and clustering, as well as classification [33], [34]. AIS is a distributed network, capable of doing parallel processing. The binary AIS is, technically, proposed according to the immune clonal selection perspective. This paper focuses on the clonal selection, which is implemented in the binary AIS. Moreover, AIS is used in the context of this work for optimizing the RBFNN-2SATRA output weight by minimizing the error of training. The AIS implementation in the RBFNN is the RBFNN-2SATRAAIS, and the involved algorithm in RBFNN-2SATRAAIS involves:

Step 1: Initialization

Initialize the population of 100 B-cells (output weights) in the system according to [35]. The representations of B-cells are as follows:

$$w_{ij} = \left(w_{1j}, w_{2j}, w_{3j}, \ldots, w_{nj}\right) \tag{19}$$

where nj is the number of weight output. The objective function for this problem involves minimizing the objective function value:

$$f_{AIS}(w_i) = \sum_{i=1}^{j} w_i \varphi_i(x) \tag{20}$$

where $w_i \in \mathbb{R}$ refers to output weight (antibody) amid Hidden Neuron in the given hidden layer, as well as the given output neuron in the given output layer. $\varphi_i(x)$ is the Gaussian Activation Function in RBFNN-2SATRA.

Step 2: Affinity Evaluation

Affinity evaluates the feasible solution value to an objective problem. This term affinity of the B-cells is the objective function in the algorithm. The affinity value of each of the solutions in the current population can be assessed as a response to applying RBFNN-2SATRA. In this work, the basic function of calculating the affinity of each solution will be in the following equation [36]:

$$Affi_i = \frac{1}{(1 + f_{AIS}(w_i))}, \quad 0 \leq fit_i \leq 1 \tag{21}$$

Based on the above equation, the lower $f_{AIS}(w_i)$ is, the higher the affinity value will be, where $f_{AIS}(w_i)$ is the objective function. When $f_{AIS}(w_i) \rightarrow \infty, Affi_i \rightarrow 0$. In contrast, $Affi_i = 1$ *if* $f_{AIS}(w_i) = 0$, therefore $Affi_i \in [0, 1]$. According to Li *et al.*, good solutions have higher affinity [36].

Step 3: Selection

Choose 50 optimum individuals (B-cells) of the population according to affinity measure to undergo the cloning operator to further diversify the population to attain better affinity.

Step 4: Cloning

This step is important because it involves cloning the best B-cells that were selected. Cloning initiates replicating the chosen B-cells utilizing a classical selection of roulette wheel to the system [14]. According to [37], β refers to the population clone number introduced by the program to search space, and consistent with [35], $\beta = 200$ is selected.

$$RC_i = \frac{Affi_i \times \beta}{\sum Affi_i}, \tag{22}$$

where RC_i is the rate of cloning or the number of clones allowed, $Affi_i$ refers to the solution's affinity value, and $\sum Affi_i$ refers to total solutions' affinity values in the population. The procedure mentioned above provides further clones of strings with lower $f_{AIS}(w_i)$ than those with higher $f_{AIS}(w_i)$, $\beta = 200$ is selected as a fixed parameter.

Step 5: Normalization

This process is important in the algorithm as a specific mechanism process, prior to the improvement by the hypermutation process. Compute the normalized affinity of the B-cells, the process is called affinity maturation. The standard formulation for B-cells normalization is given in the following equation [33]:

$$affN_i = \frac{aff_i - \min aff}{\max aff - \min aff} \tag{23}$$

where $affN_i$ is normalize the affinity of B-cells. min aff is the minima B-cells affinity value. maxaff is the maxima value of the affinity of B-cells.

Step 6: Somatic Hypermutation

Calculate the number of mutations, which is the most important step for AIS during the optimization process. This is the main event in the improved binary AIS algorithm. Also, the essential impetus of the somatic hypermutation implies improving the B-cells for a feasible solution. Based on [37], a selective pressure mechanism optimizes the B-cells' ability (output weight) so that the optimum affinity is obtained. The somatic hypermutation rate is regarded as relative to the cell affinity. The higher affinity a cell receptor possesses with antigen, the mutation rate will conversely be lower. By using this strategy, the human's immune system keeps high-affinity offspring cells, thereby confirming large mutations of the low-affinity ones to improve affinity cells [38]. The given mutation formula's number is introduced by Layeb *et al.* [35] in the following equation:

$$NM = aff\,N_i \times \frac{1}{NN} + (1 - affN_i) \times 0.01, \tag{24}$$

where NM the number of mutations, NN is the number of neurons, and $affN_i$ normalizes the affinity of B-cells (output weights). Next, generate a new solution B-cells (w_i^{new}) by using the following equation:

$$w_i^{new} = \begin{cases} rand(-5, 5), & rand(0, 1) < r \\ w_i^*, & rand(0, 1) \geq r, \end{cases} \tag{25}$$

where w_i^{new} refers to the novel B-cell when $r \in [0, 1]$. After that, the affinity of the new generation of B-cells will be calculated.

Step 7: Termination

If the termination condition is met,

$$\left| f(w_i^{new}) - y_i \right| \leq tolerance \tag{26}$$

stop and the best B-cells (optimal output weights) will be memorized; or else the algorithm will go back to the previous Step 2. The tolerance value is the termination or stopping criterion. As for the analysis, choose 0.001 to reduce statistical error [14].

VI. EXPERIMENTAL SETUP

Experimental simulation is enhanced for assessing the performance of the metaheuristic algorithms in training RBFNN in performing 2SATRA. In this paper, a real-life data set of the Malaysian monthly prices of palm oil with 7 attributes that are exerted in the established 2SATRA logic mining. This study uses a multivariate approach, in which, several new explanatory variables (attributes) are included in forecasting the movement of the Malaysian monthly price of palm oil. The attributes' price ranged from Jan 2016 to March 2020 for palm oil manufacturing. The information about the data set is presented in TABLE 1. Based on the Malaysian Palm Oil Price dataset, 60% of the data point in datasets is used for training; 40% is utilized for testing. This study utilized

60%:40% for the data set entries in training and testing for constructing an optimal learning network Tthe network in the training phase requires more entries compared with the testing phase. Otherwise, the 60%:40% ratio has a good agreement with Zamri *et al.* [15] and Kho *et al.* [18]. k Mean clustering [15] has been utilized for converting the given dataset into specific binary representation. The k-Means method, which was developed by MacQueen [39], is one of the most widely used non-hierarchical methods. It is a partitioning method, which is particularly suitable for large amounts of data. First, an initial partition with k clusters (given number of clusters) is created. Then, starting with the first object in the first cluster, Euclidean distances of all objects to all cluster foci are calculated. If an object is detected whose distance to the center of gravity of the own cluster is greater than the distance to the center of gravity (centroid) of another cluster, this object is shifted to the other cluster. Finally, the centroids of the two changed clusters are calculated again, since the compositions have changed here. These steps are repeated until each object is located in a cluster with the smallest distance to its centroid. The entire 2SATRA models have been used in Microsoft Visual C++ software using Microsoft Windows 7, in 64-bit, with 3.40 GHz processor and 4096 MB RAM, in addition to 500 GB of hard-drive specifications.

The use of C++ aims at helping users to take full command over the management of memory. It can be observed that the entire simulations have been performed in a similar device so that no possible biases can occur. Regarding the selection of the activation function, Gaussian activation function has been employed because of the association properties of every radial unit as center, as well as width. The remaining activation functions like hyperbolic activation function [40] and bipolar activation function [41], as well as Mc Culloch Pitts Activation Function [41], do not suit the presented RBFNN due to the given non-compatible classification interval. The activation function utilization has led to the RBFNN overfitting nature. Thus, the classification result has used a similar tolerance value, i.e., $Tol = 0.001$ presented by Sathasivam [42]. Selecting Tol value aims at ensuring that a possible statistical error is reduced between Target Output, as well as the output. TABLE 2 provides a summary of the parameters' lists in each of the RBFNN-2SATRA models.

VII. RESULTS AND DISCUSSION

Based on the results of the conducted experiments, the performance of the given training algorithms is measured in accordance with varied Hidden Neurons' number $4 \leq NH \leq 48$. Six different measurements were conducted to evaluate RBFNN-2SATRA models via metaheuristic algorithms, including Accuracy and Schwarz Bayesian Criterion (SBC) to evaluate accuracy of prediction. The results showed that Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), as well as Mean Absolute Percentage Error (MAPE) and Mean Absolute Relative Error (MARE) in addition to Central Process Unit time (CPU time), exhibited the complexity of structure of the RBFNN-2SATRA network according to the rising number of the neurons as displayed by this equation:

$$RMSE = \sum_{i=1}^{n} \sqrt{\frac{1}{n} (f(w_i) - y_i)^2} \tag{27}$$

RMSE [43] represents a specific standard error estimator commonly utilized in predictions, as well as classifications. RMSE, all through the given learning phase, has been measuring the error's deviation amid the existing value $f(w_i)$, as well as y_i regarding the mean $\bar{f}$. When RMSE is lower, this implies achieving better accuracy for the introduced model.

$$\text{MAE} = \sum_{i=1}^{n} \frac{1}{n} |f(w_i) - y_i| \tag{28}$$

MAE represents a specific loss function error type that directly measures the difference amid expected value, as well as an existing value. MAE, all through the given learning phase, has been measuring the over-all difference amid the given current value $f(w_i)$, as well as y_i [44]. Also, a smaller MAE value denotes the best method's fitness.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{f(w_i) - y_i}{f(w_i)} \right| \tag{29}$$

MAPE [43] has evaluated the error's size in terms of percentage. MAPE, all through the given learning phase, has been measuring the percentage difference amid the existing current value $f(w_i)$, as well as y_i. A lower MAPE resulted in better accuracy regarding the model's percentage.

$$\text{MARE} = \sum_{i=1}^{n} \frac{1}{n} \frac{|f_{\max} - f_i|}{|f_i|} \tag{30}$$

The mean absolute relative error (MARE) [45] is expressed as the average absolute value of relative differences between the current value $f(w_i)$, as well as y_i.

$$SBC = n \ln\left(\frac{\sum_{i=1}^{n} (f(w_i) - y_i)^2}{n}\right) + pa \ln(n) \tag{31}$$

where pa refers to the center's number, widths, as well as output weights. Concerning the values of SBC, lower values designate the better values [16].

On the other hand, if CPU time is lower, the model's efficiency will be enhanced. The CPU time equation is [42], [46]:

$$CPU\ time = Treaning\ Time + Testing\ Time \tag{32}$$

The CPU time represents the required time for the RBFNN-2SAT models to complete a single execution. It involves the ability, as well as the stability of the RBFNN-2SAT models.

During the testing phase, different benchmarks were used to reflect the forecasting performance of the proposed model as follows:

TABLE 2. Optimal parameters in RBFNN-2SATRA models.

AIS		PSO		GA	
Parameter	Value	Parameter	Value	Parameter	Value
Number of iteration	10000	Ω	0.6	Number of iteration	10000
β	200	ε_1	2	Selection type	Wheel selection
Population size	100	ε_2	2	Number of individuals	50
r	$[0,1]$	$rand_1 = rand_2$	$[0,1]$	Mutation ratio	1
		Number of iteration	10000	Mutation type	Uniform
				Crossover ratio	1

Testing_RMSE measures the deviation of the error between the P_i^B Number of the correct induced logic, P_i^{test} Total number of testing data.

$$\text{TESTING\_RMSE} = \sum_{i=1}^{n} \sqrt{\frac{1}{n}\left(P_i^{test} - P_i^B\right)^2} \tag{33}$$

Note that the lower value of *Testing_RMSE* signifies the small error deviation of the proposed P_i^B for the P_i^{test}.

Testing_MAE measures the deviation of the error between the P_i^B Number of the correct induced logic, P_i^{test} Total number of testing data.

$$\text{TESTING\_MAE} = \sum_{i=1}^{n} \frac{1}{n}\left|P_i^{test} - P_{i\ i}^B\right| \tag{34}$$

Note that the lower value of *Testing_MAE* signifies the small error deviation of the proposed P_i^B for the P_i^{test}.

Testing_MAPE measures the percentage of the error between the P_i^B Number of the correct induced logic, P_i^{test} Total number of testing data.

$$\text{TESTING\_MAPE} = \sum_{i=1}^{n} \frac{100}{n} \frac{\left|P_i^B - P_i^{test}\right|}{\left|P_i^{test}\right|} \tag{35}$$

Note that the lower value *Testing_MAPE* signifies the small error percentage of the proposed P_i^B concerning the P_i^{test}.

Testing_MARE measures the average absolute value of relative differences between the P_i^B Number of the correct induced logic, P_i^{test} Total number of testing data.

$$\text{TESTING\_MARE} = \sum_{i=1}^{n} \frac{1}{n} \frac{\left|P_i^B - P_i^{test}\right|}{\left|P_i^{test}\right|} \tag{36}$$

Note that the lower value of *Testing_MARE* signifies the small error average absolute value of the proposed P_i^B concerning the P_i^{test}.

Thus, when the error value is small, this shows better accuracy, which is defined as in the following:

$$Accuracy = \frac{Number\ of\ the\ correct\ induced\ logic}{Total\ number\ of\ testing\ data} \times 100\% \tag{37}$$

Accuracy determines the system's capability of training the Malaysian Palm Oil Price dataset.

The results of the three models called RBFNN-2SATRAGA, RBFNN-2SATRAPSO, and RBFNN-2SATRAAIS are summarized in FIGURE 3 to FIGURE 8 and TABLE 3 and TABLE 4. These three models are for developing the logical rule to explore the relationship among the candidate's attributes, which contributes to proper anticipation regarding the fluctuations in the price of the Malaysia Palm Oil over time. It also contributes to enhancing the export performance of the Malaysian palm oil, in addition to forecasting the spot price in the future. Based on the experimental results, the best model is RBFNN-2SATRAAIS, which can classify data in accordance with the 2SATRA logical rule using the minimum values of RMSE, MAPE, MAE, MARE, SBC, as well as CPU time. RBFNN-2SATRAAIS model showed the best performance regarding errors as the given hidden neuron number was improved as shown in FIGURE 3 to FIGURE 6. The significant features of the model, including variation and recognition, as well as the memory, in addition to learning and distributed perception, along with self-organizing affected the capability of performance.

FIGURE 3 displays the performance of RMSE in the training phase for RBFNN-2SATRAAIS, RBFNN-2SATRAPSO, and RBFNN-2SATRAGA. Relatively, the RBFNN-2SATRAAIS approach is significantly better than RBFNN-2SATRAPSO and RBFNN-2SATRAGA. This is due to the robust operators such as the cloning and somatic hypermutation feature of AIS in the training phase, which enhanced the affinity of the solutions. According to the value of MAPE in FIGURE 4, it was obvious that RBFNN-2SATRAAIS recorded MAPE consistently less value for all the numbers of hidden neurons. On the contrary, the MAPE for RBFNN-2SATRAPSO and RBFNN-2SATRAGA increased significantly until the last simulations. The lower value of MAPE has provided profound evidence of the performance of AIS in working well with RBFNN-2SATRA. The optimization operators as the somatic hypermutation in RBFNN-2SATRAAIS during obtaining the final outputs have minimized the percentage of the suboptimal state as opposed to RBFNN-2SATRAPSO and RBFNN-2SATRAGA. As for FIGURE 5, a similar result was achieved for the MAE values in comparison with RBFNN-2SATRAAIS, RBFNN-2SATRAPSO, and RBFNN-2SATRAGA. MAE is denoted as an actual error, widely used by researchers to investigate the correctness

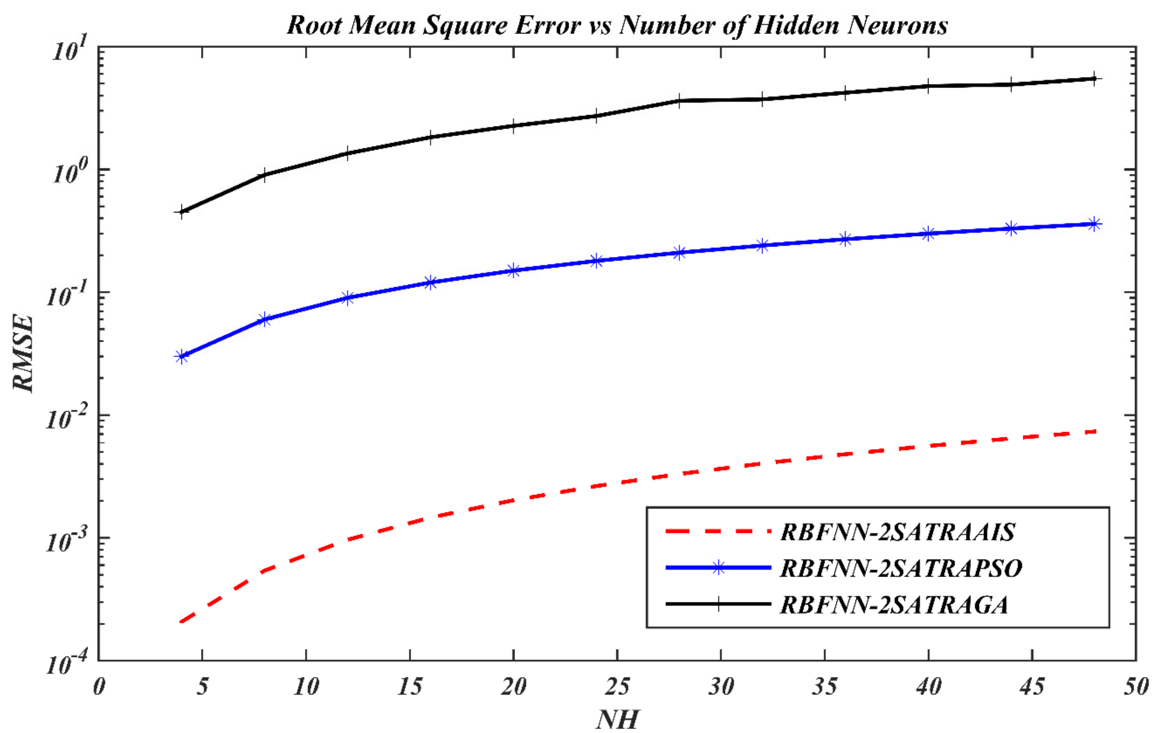

FIGURE 3. RMSE evaluation for forecasting palm oil price models.

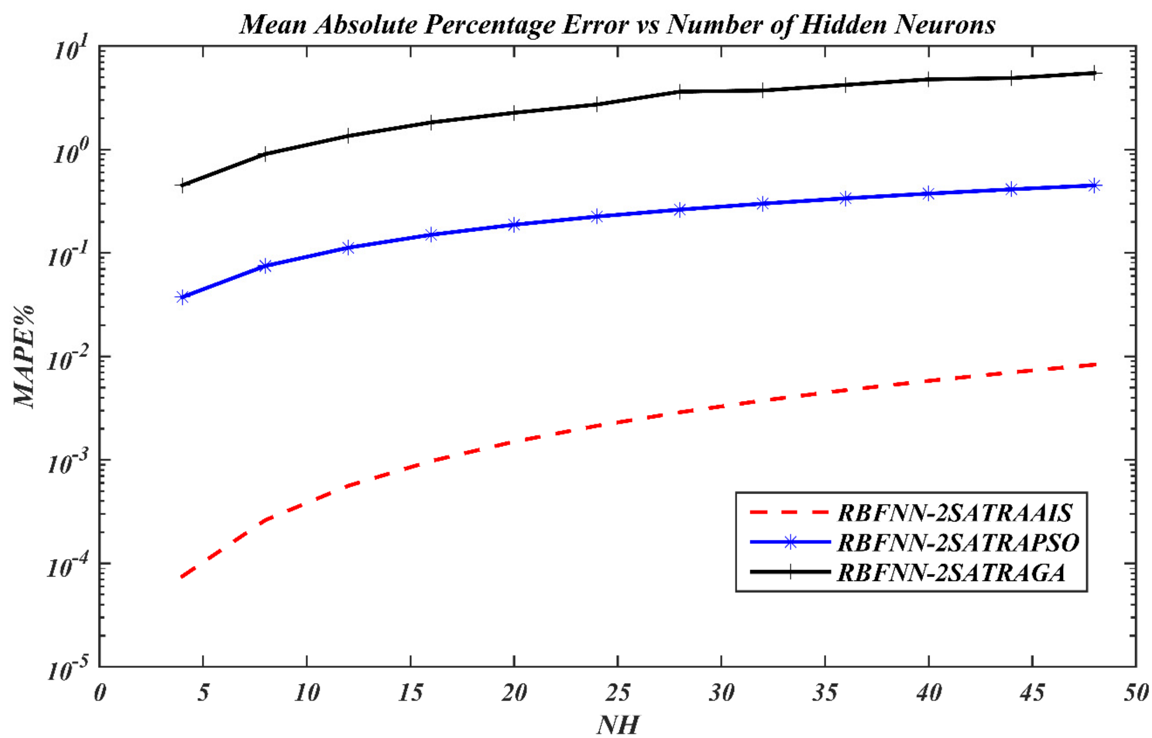

FIGURE 4. MAPE evaluation for forecasting palm oil price models.

of the solution. By observing FIGURE 5, it can be concluded that RBFNN-2SATRAAIS has a sturdier capability to train the Malaysian palm oil data set compared to RBFNN-2SATRAPSO and RBFNN-2SATRAGA due to the lower MAE values. MARE is one of the prominent error estimator tools. Based on FIGURE 6, it can be observed that RBFNN-2SATRAAIS has attained an inferior MARE value compared to RBFNN-2SATRAPSO and RBFNN-2SATRAGA.

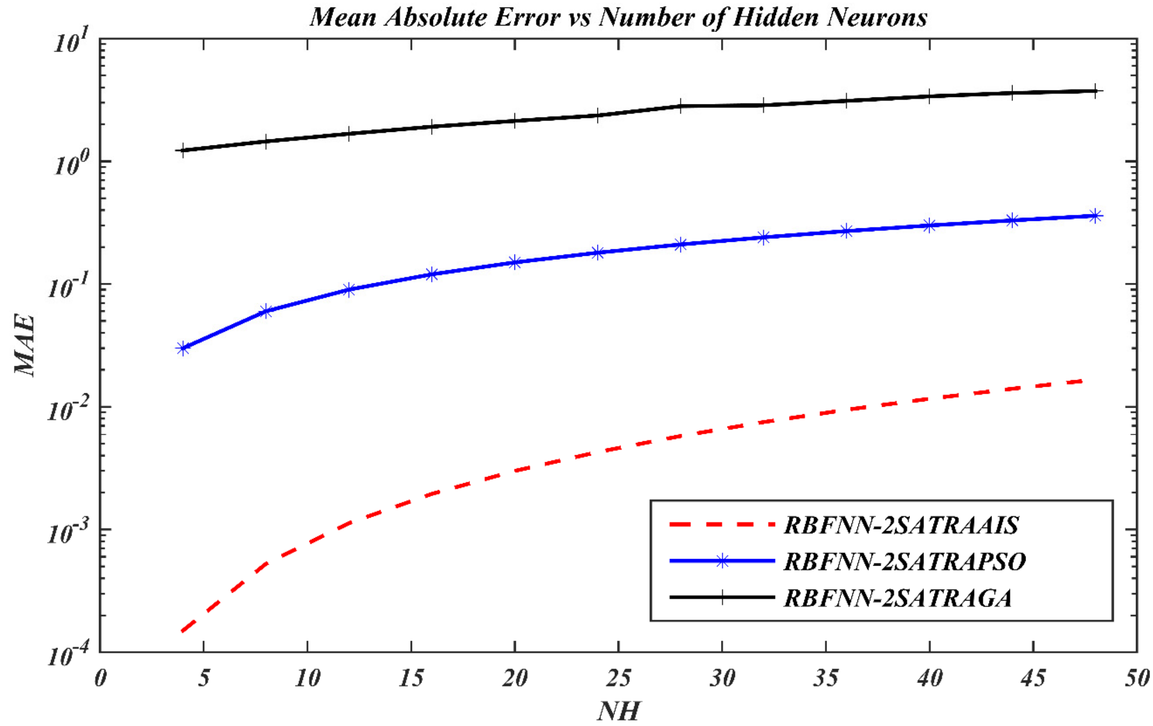

FIGURE 5. MAE evaluation for forecasting palm oil price models.

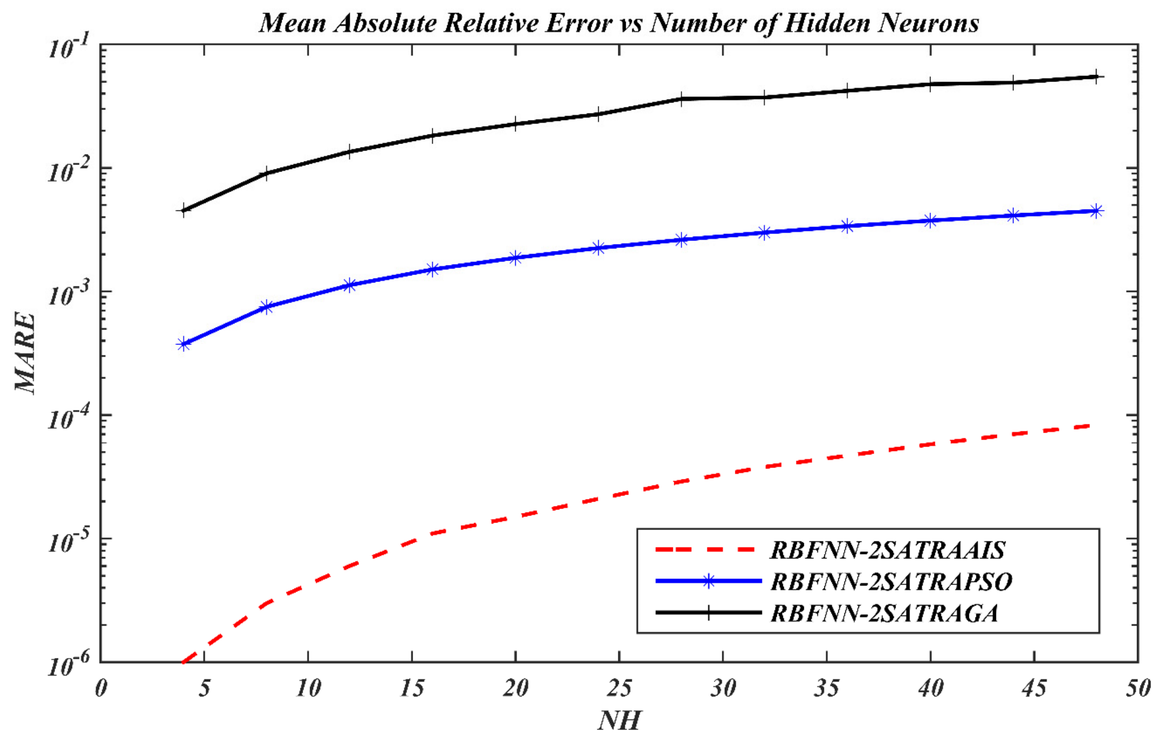

FIGURE 6. MARE evaluation for forecasting palm oil price models.

In Hamadneh *et al.* [13], the lowest SBC value showed that the model is the optimum model. The RBFNN-2SATRAAIS model showed the most optimum performance in relation to Schwarz Bayesian Criterion (SBC) even though the Hidden Neurons' number increased as shown in FIGURE 7. Concerning CPU time, the performance of RBFNN-2SATRAAIS

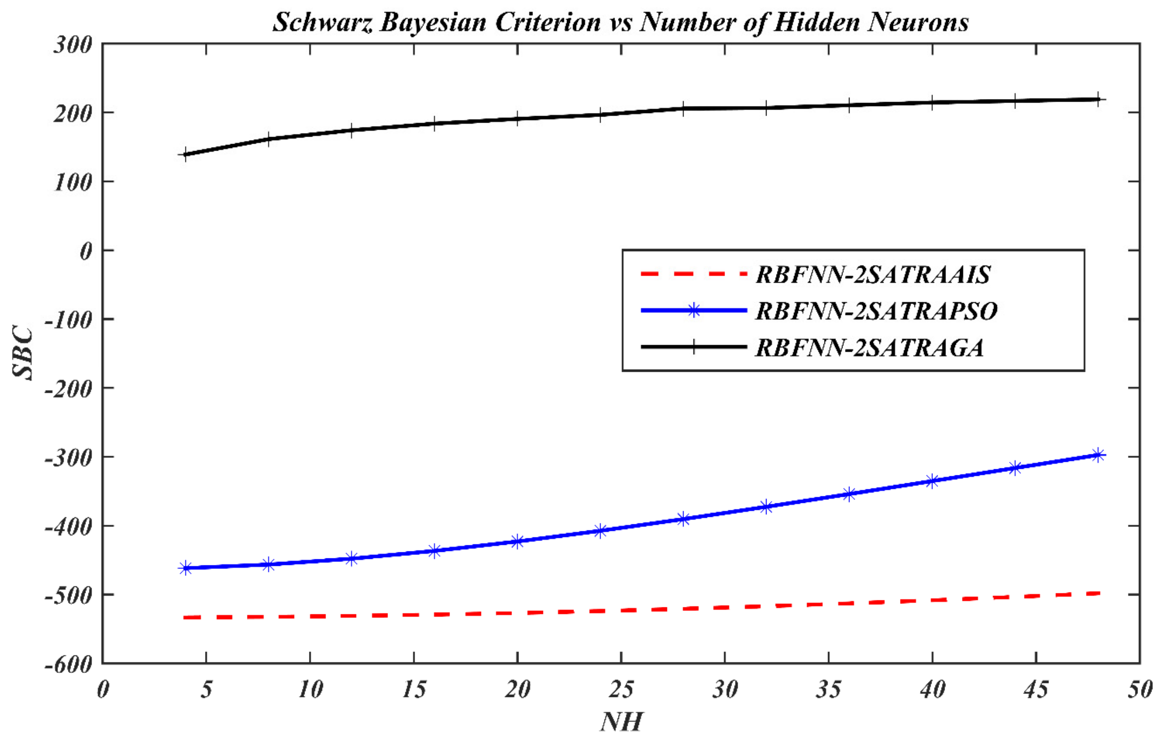

FIGURE 7. SBC evaluation for forecasting palm oil price models.

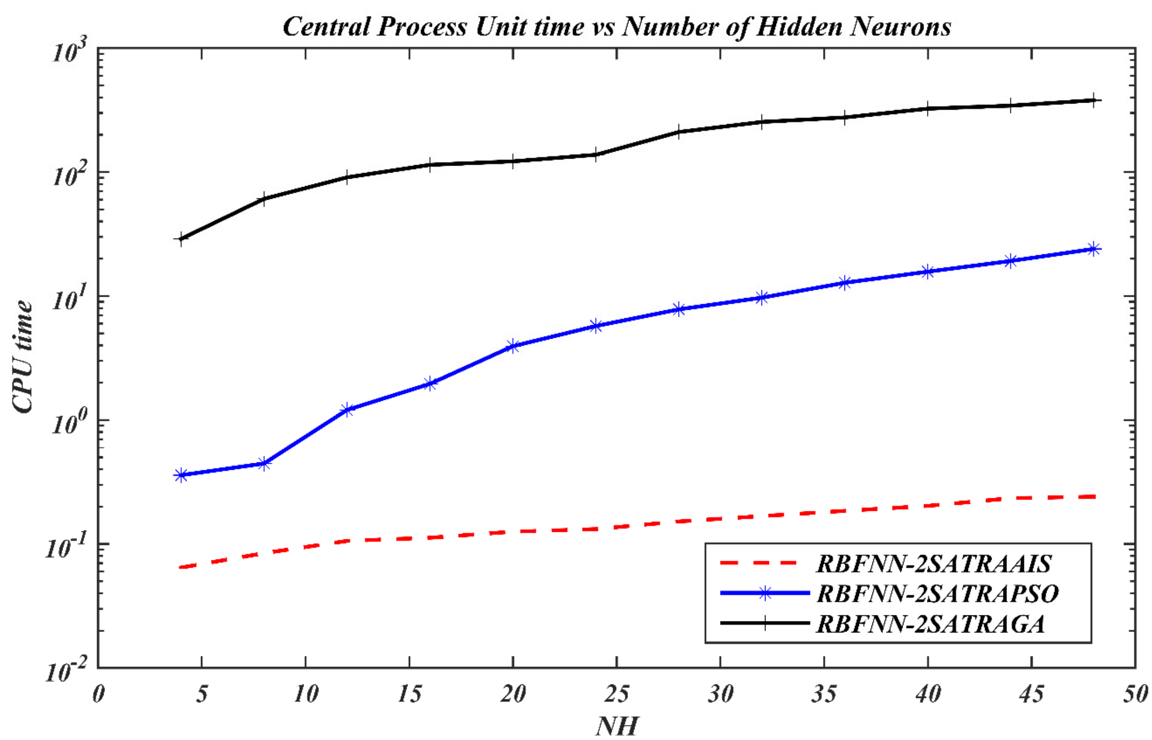

FIGURE 8. CPU time evaluation for forecasting palm oil price models.

model has been faster in comparison with the remaining models as shown in FIGURE 8. If $NH > 15$, the possibility for the RBFNN-2SATRAGA and RBFNN-2SATRAPSO of being trapped in the given state of the trial/error has increased.

TABLE 3. Testing error for all RBFNN-2SAT models.

Model	Testing RMSE	Testing MAE	Testing MAPE	Testing MARE
RBFNN-2SATGA	1.091089	0.238095	1.133787	0.011338
RBFNN-2SATPSO	0.654654	0.142857	0.680272	0.006803
RBFNN-2SATAIS	0.436436	0.095238	0.453515	0.004535

TABLE 4. Best logic and accuracy.

Models	Logic	Accuracy (%)
RBFNN-2SATRAGA	$P_{2SAT} = (A \vee B) \wedge (C \vee \neg D) \wedge (\neg E \vee \neg F)$	76.190476
RBFNN-2SATRAPSO	$P_{2SAT} = (A \vee B) \wedge (C \vee D) \wedge (\neg E \vee \neg F)$	85.714286
RBFNN-2SATRAAIS	$P_{2SAT} = (A \vee B) \wedge (C \vee D) \wedge (E \vee \neg F)$	90.476190

Trial and error made RBFNN-2SATRAGA accomplish the pre-mature convergence. However, the model of RBFNN-2SATRAGA possesses rather a greater learning error due to ineffective initial crossover. Thus, a number of iterations are required for the model of RBFNN-2SATRAG to generate output weight, which is high in quality, amid attributes, throughout which, the mutation is the single active operator. This problem has become worse in case the suboptimal output weight among attributes involves the floating number. In RBFNN-2SATRAGA, novel generations are formed via reproduction. In RBFNN-2SATRAAIS, however, novel generations are formed via cloning. In RBFNN-2SATRAAIS, the search agents' number is not constant as it increased because of the cloning operations. However, in the model of RBFNN-2SATRAGA, the search agents remained constant. In RBFNN-2SATRAAIS, the clone itself moved to the neighboring nodes. In RBFNN-2SATRAGA, a search field involves the entire population. From another perspective, the RBFNN-2SATRAPSO model has rather a lower learning error in comparison with RBFNN-2SATRAGA because of the utilization of the particle in the algorithm, thereby mimicking the introduced RBFNN-2SATRAAIS model in this work. The result of the RBFNN-2SATRAPSO model seems quite promising. However, the aspect of the control of the efficient local search is lacking [44]. The search space, in this case, for each of the particles, magnifies indefinitely, thereby leading to the suboptimal output weight between the attributes. Therefore, the RBFNN-2SATRAPSO model will converge prematurely.

TABLE 3 illustrates the respective testing error recorded for all models during testing the RBFNN-2SATRA. The testing RMSE, MAE, MAPE, and MARE recorded for RBFNN-2SATRAAIS, RBFNN-2SATRAPSO, and RBFNN-2SATRAGA were consistently similar for each $NH = 4$ until $NH = 48$. This highlights the capability of the proposed network, i.e., *2* SATRA and RBFNN in generating the best logic during the training phase, which resulted in a very minimum error during the testing phase. The training mechanism in 2SATRA in extracting the best logic to map the relationship of the attributes in RBFNN with AIS is acceptable according to the performance evaluation metrics recorded during simulation. The simulation results in TABLE 4 proved that RBFNN-2SATRAAIS is the best model - it enjoys the capability of classifying a higher accuracy percentage of (90.476190%) of the test samples compared to the findings of the RBFNN-2SATRAPSO model (85.71%) and RBFNN-2SATRAGA model (76.19%). The results showed that the RBFNN-2SATRAAIS model has the potential of obtaining the logical rule, which classifies the relationship among the candidate's attributes. This contributes to achieving a higher level of proper anticipation of the price fluctuations of Malaysia's Palm Oil. Based on the results using the RBFNN 2SATRAAIS model, the long-run relationship among the variables has been confirmed, which revealed that all independent variables play a significant role in determining the movement or behavior of palm oil price. The results also showed that the most significant variables, which affect the palm oil price, include the total exports of palm oil, the total imports of palm oil, the production of palm oil, end-stocks of palm oil, and Malaysia's real effective exchange rate. This finding showed that palm oil has an advantage in maintaining the stability of the Malaysian currency exchange. Therefore, the government and policymakers may use this information to develop the palm oil industry in Malaysia. For forecasting, the RBFNN-2SATRAAIS model seems to be the best model used in forecasting the price of Malaysian palm oil. The price fluctuations of palm oil may help policymakers modify their budgetary plans for more investment grants and incentives in the palm oil market.

VIII. CONCLUSION

The primary aim of the study involves predicting the Malaysian's monthly palm oil price. Three different models were utilized in this study to investigate the determinants of Malaysia's palm oil prices and forecast the crude palm oil price in the future. Based on the assessment, the RBFNN-2SATRAAIS model is one of the finest logic mining methods to predict the monthly price for palm oil in Malaysia. The accuracy of RBFNN-2SATRAAIS in the Malaysian palm oil data set was the highest (90.476190%), followed by RBFNN-

2SATRAPSO (85.71%) and RBFNN-SATRAGA (76.19%). According to the outcome of the experiment, the desired monthly price trend for Malaysia's palm oil has been established with the highest accuracy. Such information is useful for the government and policymakers to enhance the growth of Malaysia's palm oil industry. Therefore, it is recommended that training sessions and workshops should be held for the export department staff to encourage investment in the palm oil market.

REFERENCES

[1] N. Khalid, H. N. A. Hamidi, S. Thinagar, and N. F. Marwan, "Crude palm oil price forecasting in Malaysia: An econometric approach (Peramalan Harga Minyak Sawit Mentah Malaysia: Satu Pendekatan Ekonometrik)," *Jurnal Ekonomi Malaysia*, vol. 52, no. 3, pp. 263–278, 2018.

[2] N. Othman, Z. Yusop, and M. M. Ismail, "Identifying export competitiveness of Malaysia's processed palm oil: Using shift-share technique," *Int. J. Accounting, Finance Bus.*, vol. 3, no. 12, pp. 1–11, 2018.

[3] A. K. A. Rahman, "Impact of palm oil supply and demand on crude palm oil price behavior," Malaysia Palm Oil Board (MPOB), Kuching, Malaysia, Tech. Rep., 2012. [Online]. Available: https://www.pointers.org.my/slide/slide116.pdf

[4] M. N. Shamsudin and F. M. Arshad, "Short-term forecasting of crude palm oil prices," presented at the PORIM Int. Palm Oil Conf., Kuala Lumpur, Malaysia, Sep. 9, 1991.

[5] F. M. Arshad and R. A. Ghaffar, "Crude palm oil price forecasting: Box-Jenkins approach," *Pertanika*, vol. 9, no. 3, pp. 359–367, 1986.

[6] M. Catalano, F. Galatioto, M. Bell, A. Namdeo, and A. S. Bergantino, "Improving the prediction of air pollution peak episodes generated by urban transport networks," *Environ. Sci. Policy*, vol. 60, pp. 69–83, Jun. 2016.

[7] C. Qin, H. Zhu, T. Xu, C. Zhu, L. Jiang, E. Chen, and H. Xiong, "Enhancing person-job fit for talent recruitment: An ability-aware neural network approach," in *Proc. 41st Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Ann Arbor, MI, USA, Jun. 2018, pp. 25–34.

[8] Z. Ismail and A. Khamis, "Neural network in modeling Malaysian oil palm yield," *Amer. J. Appl. Sci.*, vol. 8, no. 8, pp. 796–803, Aug. 2011.

[9] N. Kanazawa, "Radial basis functions neural networks for nonlinear time series analysis and time-varying effects of supply shocks," *J. Macroecon.*, vol. 64, Jun. 2020, Art. no. 103210.

[10] S. Goudarzi, W. H. Hassan, A.-H. A. Hashim, S. A. Soleymani, M. H. Anisi, and O. M. Zakaria, "A novel RSSI prediction using imperialist competition algorithm (ICA), radial basis function (RBF) and firefly algorithm (FFA) in wireless networks," *PLoS ONE*, vol. 11, no. 7, Jul. 2016, Art. no. e0151355.

[11] Y.-C. Ho and D. L. Pepyne, "Simple explanation of the no free lunch theorem of optimization," *Cybern. Syst. Anal.*, vol. 38, no. 2, pp. 292–298, 2002.

[12] Z. Jing, J. Chen, and X. Li, "RBF-GA: An adaptive radial basis function metamodeling with genetic algorithm for structural reliability analysis," *Rel. Eng. Syst. Saf.*, vol. 189, pp. 42–57, Sep. 2019.

[13] N. Hamadneh, W. A. Khan, S. Sathasivam, and H. C. Ong, "Design optimization of pin fin geometry using particle swarm optimization algorithm," *PLoS ONE*, vol. 8, no. 5, May 2013, Art. no. e66080.

[14] M. A. Mansor, M. S. M. Kasihmuddin, and S. Sathasivam, "Modified artificial immune system algorithm with Elliot Hopfield neural network for 3-satisfiability programming," *J. Informat. Math. Sci.*, vol. 11, no. 1, pp. 81–98, 2019.

[15] N. E. Zamri, M. A. Mansor, M. S. M. Kasihmuddin, A. Alway, S. Z. M. Jamaludin, and S. A. Alzaeemi, "Amazon employees resources access data extraction via clonal selection algorithm and logic mining approach," *Entropy*, vol. 22, no. 6, p. 596, May 2020.

[16] S. A. Alzaeemi and S. Sathasivam, "Artificial immune system in doing 2-satisfiability based reverse analysis method via a radial basis function neural network," *Processes*, vol. 8, no. 10, p. 1295, Oct. 2020.

[17] S. Sathasivam and W. A. T. Wan Abdullah, "Logic mining in neural network: Reverse analysis method," *Computing*, vol. 91, no. 2, pp. 119–133, Feb. 2011.

[18] L. C. Kho, M. S. M. Kasihmuddin, M. Mansor, and S. Sathasivam, "Logic mining in league of legends," *Pertanika J. Sci. Technol.*, vol. 28, no. 1, pp. 211–225, Jan. 2020.

[19] R. Sikora and S. Piramuthu, "Efficient genetic algorithm based data mining using feature selection with hausdorff distance," *Inf. Technol. Manage.*, vol. 6, no. 4, pp. 315–331, Oct. 2005.

[20] J. Yuan, "An anomaly data mining method for mass sensor networks using improved PSO algorithm based on spark parallel framework," *J. Grid Comput.*, vol. 18, no. 2, pp. 251–261, Jun. 2020.

[21] S. Sathasivam, "Application of neural networks in predictive data mining," *SSJ*, vol. 1, no. 2, p. 2, 2011.

[22] S. S. Sathasivam, S. A. Alzaeemi, M. T. Ismail, and V. VaniPachala, "Palm oil price forecasting in Malaysia using 2 satisfiability based reverse analysis method via radial basis function neural network," *Solid State Technol.*, vol. 63, no. 1s, pp. 334–339, 2020.

[23] D. Lowe, "Adaptive radial basis function nonlinearities, and the problem of generalisation," in *Proc. 1st IEE Int. Conf. Artif. Neural Netw.*, Oct. 1989, vol. 1, no. 313, pp. 171–175.

[24] J. Moody and C. J. Darken, "Fast learning in networks of locally-tuned processing units," *Neural Comput.*, vol. 1, no. 2, pp. 281–294, Jun. 1989.

[25] S.-B. Roh, S.-K. Oh, W. Pedrycz, K. Seo, and Z. Fu, "Design methodology for radial basis function neural networks classifier based on locally linear reconstruction and conditional fuzzy C-Means clustering," *Int. J. Approx. Reasoning*, vol. 106, pp. 228–243, Mar. 2019.

[26] N. Hamadneh, S. Sathasivam, S. L. Tilahun, and O. H. Choon, "Learning logic programming in radial basis function network via genetic algorithm," *J. Appl. Sci.*, vol. 12, no. 9, pp. 840–847, Apr. 2012.

[27] F. Friedrichs and M. Schmitt, "On the power of Boolean computations in generalized RBF neural networks," *Neurocomputing*, vol. 63, pp. 483–498, Jan. 2005.

[28] M. Awad, "Optimization RBFNNs parameters using genetic algorithms: Applied on function approximation," *Int. J. Comput. Sci. Secur.*, vol. 4, no. 3, pp. 295–307, 2010.

[29] L. J. Eshelman and J. D. Schaffer, "Real-coded genetic algorithms and interval-schemata," in *Foundations of Genetic Algorithms*, vol. 2. Amsterdam, The Netherlands: Elsevier, 1993, pp. 187–202.

[30] S. Alzaeemi, M. A. Mansor, M. K. M. Kasihmuddin, S. Sathasivam, and M. Mamat, "Radial basis function neural network for 2 satisfiability programming," *Indonesian J. Elect. Eng. Comput. Sci.*, vol. 18, no. 1, pp. 459–469, 2020.

[31] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. MHS95. Proc. 6th Int. Symp. Micro Mach. Hum. Sci.*, 1995, October, pp. 39–43.

[32] Y. Fukuyama and H. Yoshida, "A particle swarm optimization for reactive power and voltage control in electric power systems," in *Proc. Congr. Evol. Comput.*, vol. 1, May 2001, pp. 87–93.

[33] L. N. De Castro and F. J. Von Zuben, "Artificial immune systems: Part II—A survey of applications," FEEC/Univ. Campinas, Campinas, Brazil, Tech. Rep., 2000. [Online]. Available: https://neuro.bstu.by/ai/To-dom/My_research/failed%201%20subitem/For-research/D-mining/Anomaly-D/Immune/Litle-relevant/rtdca0200.pdf

[34] A. Layeb, "A clonal selection algorithm based tabu search for satisfiability problems," *J. Adv. Inf. Technol.*, vol. 3, no. 2, pp. 138–146, May 2012.

[35] A. Layeb, A. H. Deneche, and S. Meshoul, "A new artificial immune system for solving the maximum satisfiability problem," in *Proc. Int. Conf. Ind., Eng. Appl. Appl. Intell. Syst.*, Jun. 2010, pp. 136–142.

[36] Z. Li, G. He, M. Li, L. Ma, Q. Chen, J. Huang, J. Cao, S. Feng, H. Gao, and S. Wang, "RBF neural network based RFID indoor localization method using artificial immune system," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Jun. 2018, pp. 2837–2842.

[37] L. N. de Castro and F. J. Von Zuben, "Learning and optimization using the clonal selection principle," *IEEE Trans. Evol. Comput.*, vol. 6, no. 3, pp. 239–251, Jun. 2002.

[38] J. Timmis and M. Neal, "A resource limited artificial immune system for data analysis," in *Research and Development in Intelligent Systems XVII*. London, U.K.: Springer, 2001, pp. 19–32.

[39] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, Jun. 1967, vol. 1, no. 14, pp. 281–297.
[40] M. A. Mansor and S. Sathasivam, "Accelerating activation function for 3-satisfiability logic programming," *Int. J. Intell. Syst. Appl.*, vol. 8, no. 10, p. 44, 2016.
[41] M. A. Mansor and S. Sathasivam, "Performance analysis of activation function in higher order logic programming," in *AIP Conference Proceedings (*, vol. 1750, no. 1. Melville, NY, USA: AIP Publishing LLC, Jun. 2016, Art. no. 030007.
[42] S. Sathasivam, "Upgrading logic programming in Hopfield network," *Sains Malaysiana*, vol. 39, no. 1, pp. 115–118, 2010.
[43] M. A. Mansor, M. Jamaludin, M. S. M. Kasihmuddin, S. A. Alzaeemi, M. F. M. Basir, and S. Sathasivam, "Systematic Boolean satisfiability programming in radial basis function neural network," *Processes*, vol. 8, no. 2, pp. 214–229, 2020.
[44] M. S. B. M. Kasihmuddin, M. A. B. Mansor, S. A. Alzaeemi, and S. Sathasivam, "Satisfiability logic analysis via radial basis function neural network with artificial bee colony algorithm," *Int. J. Interact. Multimedia Artif. Intell.*, to be published, doi: 10.9781/ijimai.2020.06.002.
[45] M. Despotovic, V. Nedic, D. Despotovic, and S. Cvetanovic, "Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation," *Renew. Sustain. Energy Rev.*, vol. 56, pp. 246–260, Apr. 2016.
[46] S. A. Alzaeemi, S. Sathasivam, and M. Shareduwan, "Maximum 2-satisfiability in radial basis function neural network," *J. Qual. Meas. Anal.*, vol. 16, no. 1, pp. 107–115, 2020.

SHEHAB ABDULHABIB SAEED ALZAEEMI received the bachelor's degree of education (science) from Taiz Universiti, in 2004, and the M.Sc. degree in mathematics from Universiti Sains Malaysia, in 2016, where he is currently pursuing the Ph.D. degree. He has several publications in the neural network. His research interests mainly focus on neural network, logic programming, and data mining. He was a Fellow under the Academic Staff Training System of Sana'a Community College from 2005 to 2014.

SARATHA SATHASIVAM received the M.Sc. and B.Sc. (Ed) degrees from Universiti Sains Malaysia, and the Ph.D. degree from Universiti Malaya, Malaysia. she is currently an Associate Professor with the School of Mathematical Sciences, Universiti Sains Malaysia. She has published widely in journals and proceedings and has collaboration with researchers from different countries. Her research interests are neural networks, agent-based modeling, data mining, and constrained optimization problem.