Forecasting Energy Consumption Using a Novel Hybrid Dipper Throated Optimization and Stochastic Fractal Search Algorithm

Doaa Sami Khafaga[1], El-Sayed M. El-kenawy[2], Amel Ali Alhussan[1,*] and Marwa M. Eid[3]

[1]Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P. O. Box 84428, Riyadh, 11671, Saudi Arabia
[2]Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura, 35111, Egypt
[3]Faculty of Artificial Intelligence, Delta University for Science and Technology, Mansoura, Egypt
*Corresponding Author: Amel Ali Alhussan. Email: AAAlHussan@pnu.edu.sa

Abstract: The accurate prediction of energy consumption has effective role in decision making and risk management for individuals and governments. Meanwhile, the accurate prediction can be realized using the recent advances in machine learning and predictive models. This research proposes a novel approach for energy consumption forecasting based on a new optimization algorithm and a new forecasting model consisting of a set of long short-term memory (LSTM) units. The proposed optimization algorithm is used to optimize the parameters of the LSTM-based model to boost its forecasting accuracy. This optimization algorithm is based on the recently emerged dipper-throated optimization (DTO) and stochastic fractal search (SFS) algorithm and is referred to as dynamic DTOSFS. To prove the effectiveness and superiority of the proposed approach, five standard benchmark algorithms, namely, stochastic fractal search (SFS), dipper throated optimization (DTO), whale optimization algorithm (WOA), particle swarm optimization (PSO), and grey wolf optimization (GWO), are used to optimize the parameters of the LSTM-based model, and the results are compared with that of the proposed approach. Experimental results show that the proposed DDTOSFS + LSTM can accurately forecast the energy consumption with root mean square error RMSE of 0.00013, which is the best among the recorded results of the other methods. In addition, statistical experiments are conducted to prove the statistical difference of the proposed model. The results of these tests confirmed the expected outcomes.

Keywords: Stochastic fractal search; dipper throated optimization; energy consumption; long short-term memory; prediction models

1 Introduction

The consumption of energy is increased over time and thus increases the consumption of fossil fuels, which might lead to the destruction of our ecosystem irreversibly. Therefore, we need to find efficient and more innovative technological methods for decreasing energy consumption in all sectors to protect our environment and ensure future energy supply. Furthermore, these technologies may assist in making more intelligent decisions about the energy consumption of all types of equipment more effectively.

The development of smart households is a cost-effective approach that can help in saving energy. Smart households are usually based on a set of relays, controllers, sensors, and meters, in addition to a control system to manage the full operation and interaction between them. The main goal of a smart household is to collect relevant information about the environment and behavior of residents along with their habits, to make efficient decisions regarding the reduction of energy consumption. A smart household can learn the routines and lifestyles of its residents using machine learning techniques, allowing activities and operations to be automated. Based on the absence and presence of residents, temperature, and light, for example, can be automatically regulated. In addition, the systems of smart households can be connected with switchboards and other traditional systems to optimize energy consumption [1]. For linking virtual and physical devices in a smart household to the Internet, the Internet of Things (IoT) can be utilized. The main advantage of IoT is that it enables users to automate the functioning of their devices, collect information and remotely manage them. The scales of applicability of IoT are various, starting from simple appliances, such as washing machines and refrigerators in a household, on a small scale, to the transportation system of a city in complex infrastructure, on a large scale.

A significant reduction of environmental effects due to saving can be achieved by predicting the potential energy consumption of a household [2–4]. Consequently, a lot of time and effort has been invested in developing solutions to energy consumption optimization in a smart household. The most essential of these solutions is to use the least amount of energy possible while retaining customer satisfaction [5]. Several machine learning (ML) techniques for energy prediction of appliances in smart households have been suggested for this purpose [6,7]. Traditional machine learning methods such as super vector machines, k-nearest neighbor, neural networks, multi-layer perceptron, decision trees, deep learning, and ensemble approaches are extensively used to provide efficient solutions to this problem of accurate energy consumption prediction.

In this paper, a novel approach is proposed for predicting energy consumption using recent advances in machine learning. The prediction of the energy consumption is performed using a new LSTM-based model. To improve the performance of the proposed LSTM-based model, we developed a new optimization algorithm based on the DTO and SFS optimization algorithms to adjust the parameters of the LSTM-based model. To show the effectiveness of the proposed approach, we employed a publicly available dataset that contains 121273 measurements gathered over ten years in the United States. The experiments based on the given dataset show that the proposed approach outperforms the other forecasting approaches considered in this work.

The following organization is followed in presenting the contents of this work. The relevant related works are presented in Section 2. The proposed methodology for energy consumption is explained in Section 3. The results of the conducted experiments are discussed in Section 4. The conclusion is then presented in Section 5.

2 Related Works

In recent research, a variety of techniques to energy consumption prediction have been developed, with the goal of developing effective energy consumption prediction systems to enhance energy savings and lessen environmental consequences. Most of these methods start with historical energy usage time series data and use machine learning techniques to create a forecast model. Authors in [6] used artificial neural network (ANN), linear regression (LR), and decision tree (DT) to forecast energy consumption in cottages, dwellings, and government buildings. They also looked at the impact of the number of people, family size, building time, house style, and electrical equipment. Summer and winter results were obtained. In comparison to ANN and LR, their findings demonstrate that DT has higher accuracy.

In the same way, the authors in [8] present three models for predicting energy consumption based on linear regression multiple layers (LRML) and simple linear regression (SLR). They also took into account three distinct scenarios. The training dataset for energy consumption was divided into three groups: annually, daily, and hourly peak hours. In comparison to the LRML, their findings reveal that the SLR-based algorithm is better at forecasting yearly and hourly-based forecasts.

In [9], the authors suggested using a graphical energy consumption prediction model. They target figuring out how energy consumption is related to human behaviors and routines, specifically the rate and duration of use of electronic household devices. Temperature differences between the outdoor and indoor environments and the energy required utilities such as heating are among the model factors.

The study's objective was to foretell energy usage patterns for the coming week and month. The extreme learning machine (ELM), the adaptive neuro-fuzzy inference system (ANFIS), and the ANN are usually introduced as methods for predicting residential energy usage. In addition, they experimented with a wide range of membership functions, both in terms of quantity and variety, to determine what made for the most effective ANFIS architecture. They have also tried out their method with different ANN architectures and levels of hidden complexity. In [10], researchers report the success of using a recurrent neural network trained on hourly energy consumption data to predict the cooling and heating energy needs of a building based on the current weather and the present time.

For estimating the yearly energy consumption of households, authors in [11] used SVM, general regression neural network (GRNN), radial basis function neural network (RBFNN), and backpropagation ANN. These models were developed using real-world data from 59 households and tested on nine others. In comparison to other models, their findings reveal that GRNN and SVM are better suited to this problem. However, a test of all approaches revealed that SVM predicted the potential consumption with more accuracy and was superior to the others.

The authors used SVM and MLP in [12] to solve the problem of predicting energy consumption. In addition, they improved the predictions of district consumption by grouping households according to their consumption patterns. Moreover, they illustrated how the normalized root mean squared error diminishes as the number of clients increases in an empirical plot. Up to 782 households are included in this research. SVM was also used to forecast monthly energy consumption in tropical regions by the authors of [13]. According to their findings based on three years of power usage data, SVM has a strong prediction performance.

The authors in [14] used clustering-based Short-Term Load prediction. The prediction is carried out for the next day's 48 half-hourly loads. The daily average load of all training and testing patterns is determined for each day, and the patterns are grouped using a threshold value between the testing pattern's daily average load and the training pattern's daily average load. Their findings reveal that

the SVM clustering-based strategy is more accurate than the results produced without clustering the input patterns. The SVM outcome for predicting was better than the Auto Regressive modeling for the same sort of forecasting, according to the authors in [15]. SVM had a root means square error (RMSE) of 0.0215 while testing data, whereas LR had an RMSE of 0.0376. The research revealed that the prediction was more accurate since the RMSE was lower.

Deep learning is a very appealing alternative for longer-term predictions because it can represent expressive functions through numerous layers of abstraction [16]. In [17], the authors used a deep recurrent neural network to predict the energy consumption needed for heating in a commercial building in the United States. They looked at the deep model's performance across a horizon of long-to-medium-term prediction.

3 The Proposed Methodology

This section discusses the overall methodology proposed to forecast energy consumption. In addition, the proposed optimization algorithm is explained.

3.1 General Scheme

The study's overarching design is shown in Fig. 1. After preparing and preprocessing the dataset, six optimization algorithms are employed to optimize the parameters of a new LSTM-based model to improve the prediction accuracy of the energy consumption. These optimization algorithms include the proposed hybrid optimization algorithm (dynamic DTO and SFS algorithms and is called DDTOSFS), the standard DTO algorithm, the standard SFS algorithm, the standard particle swarm optimization (PSO) algorithm, the standard grey wolf optimization (GWO) algorithm, and the standard whale optimization algorithm (WOA). The optimized LSTM models are: DDTOSFS + LSTM, PSO + LSTM, DTO + LSTM, SFS + LSTM, GWO + LSTM, and WOA + LSTM). The forecasting of energy consumption is made in terms of these optimized models, and the accuracy of those predictions is measured using statistical metrics.
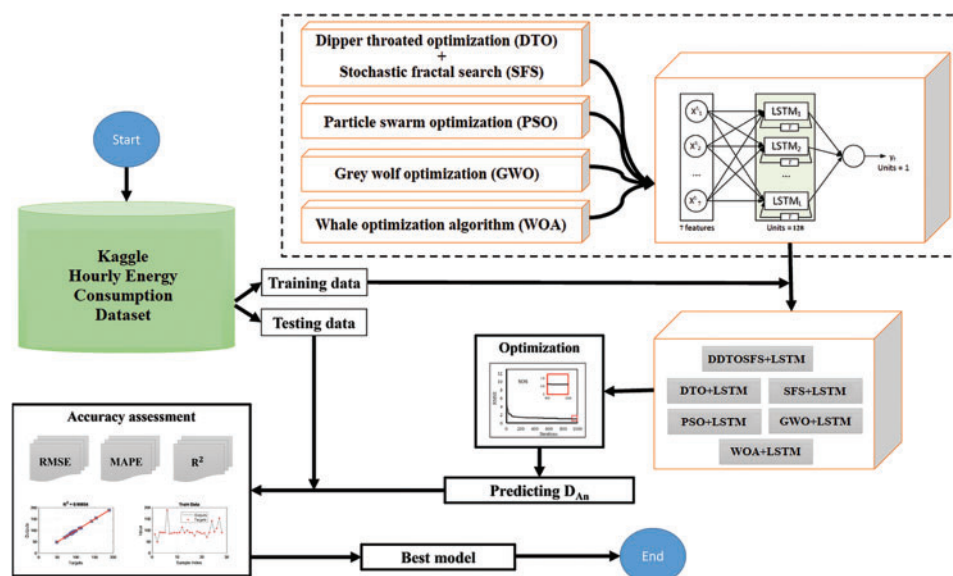

Figure 1: The overall steps of the approach proposed for forecasting energy consumption

3.2 Long Short-Term Memory (LSTM)

The LSTM model is an enhancement of the traditional neural network model that gained the capability of solving a variety of problems. Long-term memory retention is one of LSTM's key advantages that makes it suitable for modeling temporal dependencies and time sequence predictions, such as the case of IoT energy consumption prediction. Fig. 2 depicts the structure of the proposed energy consumption prediction model based on LSTM units. Each LSTM unit consists of three gates: output o, forget f, and input i. The cell state is referred to as c. The information flow in this model is usually controlled by the three gates [18]. The input at time t, is denoted by X^t. The output of the hidden layer is h^t, and its previous output is denoted by h^{t-1}. The cell input and output are referred to as $\widetilde{c}^t$, c^t, respectively, and the previous state is denoted by c^{t-1}. The measurement of c^t and h^t can be performed by calculating the following equations first.

$$i^t = \sigma\left(W_i\, X^t + U_i h^{t-1} + b_i\right) \tag{1}$$

$$f^t = \sigma\left(W_f\, X^t + U_f h^{t-1} + b_f\right) \tag{2}$$

$$o^t = \sigma\left(W_o\, X^t + U_o h^{t-1} + b_o\right) \tag{3}$$

$$\widetilde{c}^t = tanh\left(W_c\, X^t + U_c h^{t-1} + b_c\right) \tag{4}$$

where W_i, W_f, and W_o refer to matrices of weights used to connect X^t with the input cell and the three gates. On the other hand, U_i, U_f, who are matrices of weights used to connect h^{t-1} with the input cell and the three gates. The bias terms are denoted by b_i, b_f, b_o, and b_c, tanh and σ refer to the tangent and sigmoid functions, respectively. Moreover, the measurement of the cell output state is calculated as follows.

$$c^t = i^t \odot \widetilde{c}^t + f^t \odot c^{t-1} \tag{5}$$

The final output of the LSTM cell is measured using the following equation.

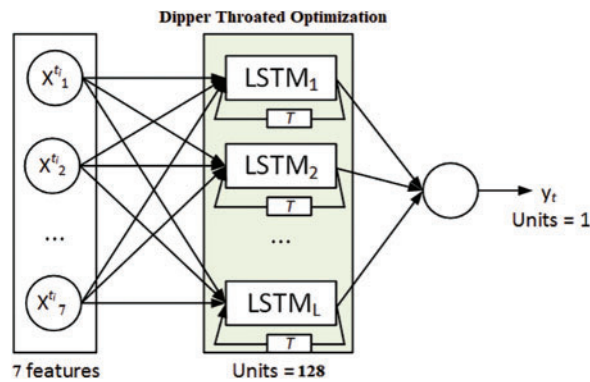$$h^t = o^t \odot tanh\ tanh\ \left(c^t\right) \tag{6}$$

Figure 2: The structure of the proposed LSTM-based forecasting model

3.3 Dipper Throated Optimization (DTO)

The DTO algorithm assumes that there are swimming and flying folks of birds looking for food. The ability of these birds to find food depends of their locations and speed while search for food. The following matrices are used to depict the birds' locations (L) and speeds (S).

$$L = \begin{bmatrix} L_{1,1} & L_{1,2} & L_{1,3} & \ldots & L_{1,d} \\ L_{2,1} & L_{2,2} & L_{2,3} & \ldots & L_{2,d} \\ L_{3,1} & L_{3,2} & L_{3,3} & \ldots & L_{3,d} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ L_{m,1} & L_{m,2} & L_{m,3} & \ldots & L_{m,d} \end{bmatrix} \tag{7}$$

$$S = \begin{bmatrix} S_{1,1} & S_{1,2} & S_{1,3} & \ldots & S_{1,d} \\ S_{2,1} & S_{2,2} & S_{2,3} & \ldots & S_{2,d} \\ S_{3,1} & S_{3,2} & S_{3,3} & \ldots & S_{3,d} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ S_{m,1} & S_{m,2} & S_{m,3} & \ldots & S_{m,d} \end{bmatrix} \tag{8}$$

where $L_{i,j}$ refers to the ith bird in the jth dimension for i $\in$ 1, 2, 3, $\ldots$, m and j $\in$ 1, 2, 3, $\ldots$, d, and its speed is indicated by $S_{i,j}$ for the jth dimension. For each bird, the following matrix represents the values of the fitness functions $h = h_1, h_2, h_3, \ldots, h_n$.

$$h = \begin{bmatrix} h_1\left(L_{1,1}, L_{1,2}, L_{1,3}, \ldots, L_{1,d}\right) \\ h_2\left(L_{2,1}, L_{2,2}, L_{2,3}, \ldots, L_{2,d}\right) \\ h_3\left(L_{3,1}, L_{3,2}, L_{3,3}, \ldots, L_{3,d}\right) \\ \ldots \\ h_m\left(L_{m,1}, L_{m,2}, L_{m,3}, \ldots, L_{m,d}\right) \end{bmatrix} \tag{9}$$

Finally, the location of swimming birds can be updated using the following formulas.

$$L(i+1) = \begin{cases} L_{best}(i) - K_1. |K_2.L_{best}(i) - L(i)| & if\ R < 0.5 \\ L(i) + S(i+1) & otherwise \end{cases} \tag{10}$$

$$S(i+1) = K_3 S(i) + K_4 r_1 (Lbest(i) - L(i)) + K_5 r_2 (LGbest - L(i)) \tag{11}$$

where $K_1, \ldots K_5, r_1$, and r_2, are adaptive variables whose values are changed during the optimization process based on a selection of random values. The DTO algorithm detailed steps are depicted in Fig. 3a [19].

3.4 Stochastic Fractal Search Algorithm

The natural phenomena of growth served as inspiration for the development of the powerful meta-heuristic algorithm known as stochastic fractal search (SFS). For optimality, this method mimics fractal characteristics. Here we discuss the Diffusion and the Update processes, the two key components of SFS's search for an optimal solution. To maximize the likelihood of discovering a better solution and to avoid local minima, the diffusion process involves each point diffusing around its present position to use the search space. By using a Gaussian walk, this method may produce an

infinite number of new points. Two statistical approaches are used in the updating process to efficiently probe the search space. The initial statistical process involves ordering all the points according to their fitness values and assigning each rank a probability value. The steps of the SFS algorithm are depicted in the flowchart shown in Fig. 3b.
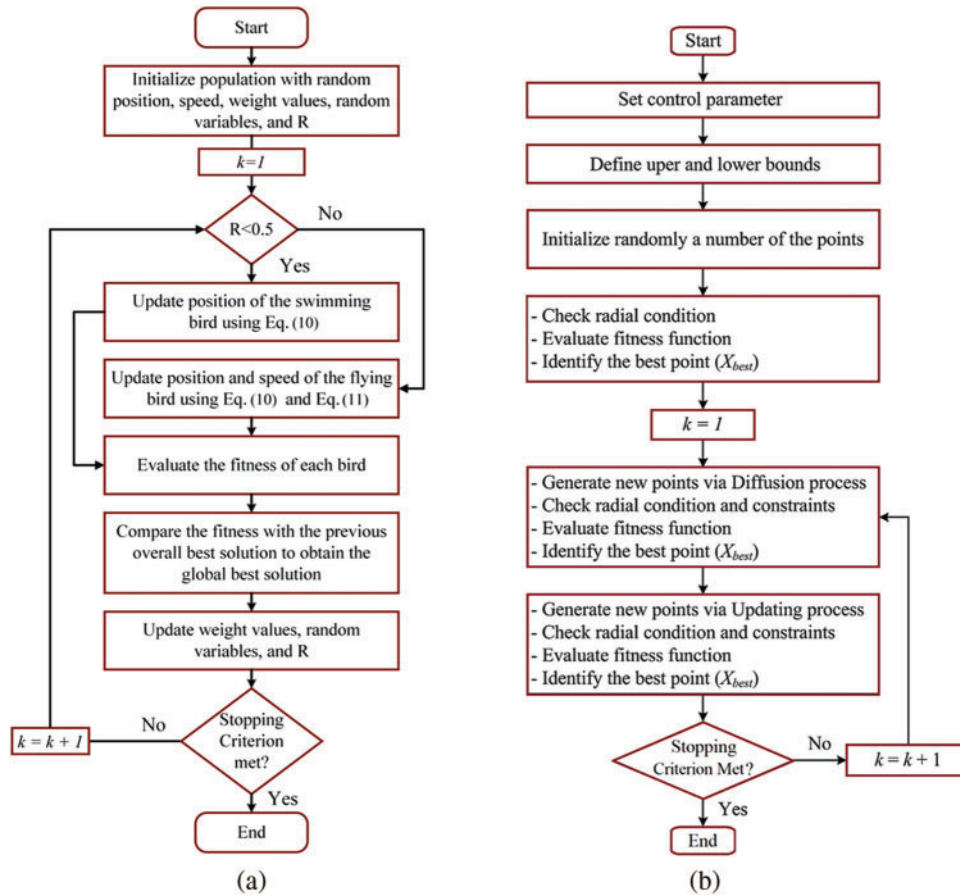

Figure 3: Flowcharts of the employed optimization algorithms (a) DTO algorithm and (b) SFS algorithm

3.5 The Proposed DDTOSFS Optimization Algorithm

The proposed optimization algorithm is based on exploring the search space using the DTO and SFS algorithms interchangeably. The interchange between the two algorithms is based on an iteration number (k); where the DTO algorithm works with the even iterations and the SFS algorithm works with the odd iterations. The steps of the proposed dynamic DTOSFS algorithm are shown in the flowchart depicted in Fig. 4.
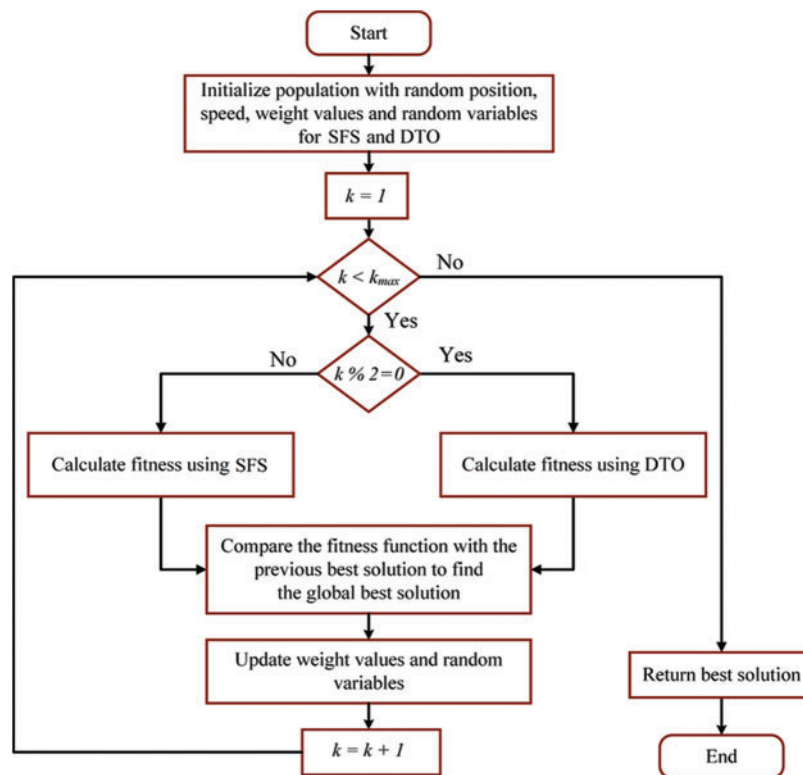
Figure 4: Flowcharts of the proposed dynamic DTOSFS algorithm

4 Experimental Results

The publicly available hourly energy consumption dataset [20] is employed to validate the proposed method's effectiveness. This dataset consists of 121273 measurements gathered over more than 10 years. To use this dataset, a set of preprocessing steps are performed to ensure the integrity and consistency of the dataset samples. These preprocessing steps include (Data resampling using various parameter optimization methods and various base models, Data normalization, Scaling the values between [0, 1], Handling outliers (using standard deviation), Handling missing values, and Data Smoothing (exponential smoothing)). Fig. 5 shows the visualization of the energy consumption distribution during over 24 h/day of a sample day in the dataset as shown in Fig. 5a, over 12 months of a sample year as shown in Fig. 5b, over the full set of years as shown in Fig. 5c, and the histogram of the energy consumption values as shown in Fig. 5d.

After data preprocessing, the dataset is divided into testing (20%) and training (80%). The LSTM's parameters are optimized using the proposed DDTOSFS algorithm, which is trained using the data from the training set. The following conditions are chosen for the training procedure. A total of 30 populations will be used, with a maximum of 20 iterations, and 20 runs will be performed. Additionally, five basic models are trained using the same training data to determine which model will benefit most from optimization. Multi-layer perceptron (MLP), Support vector regression (SVR), K-nearest neighbor (KNN), random forest (RF), and the regular LSTM are all examples of such benchmark forecasting models [21,22]. The eight evaluation metrics used to measure the performance of these models are presented in Table 1.
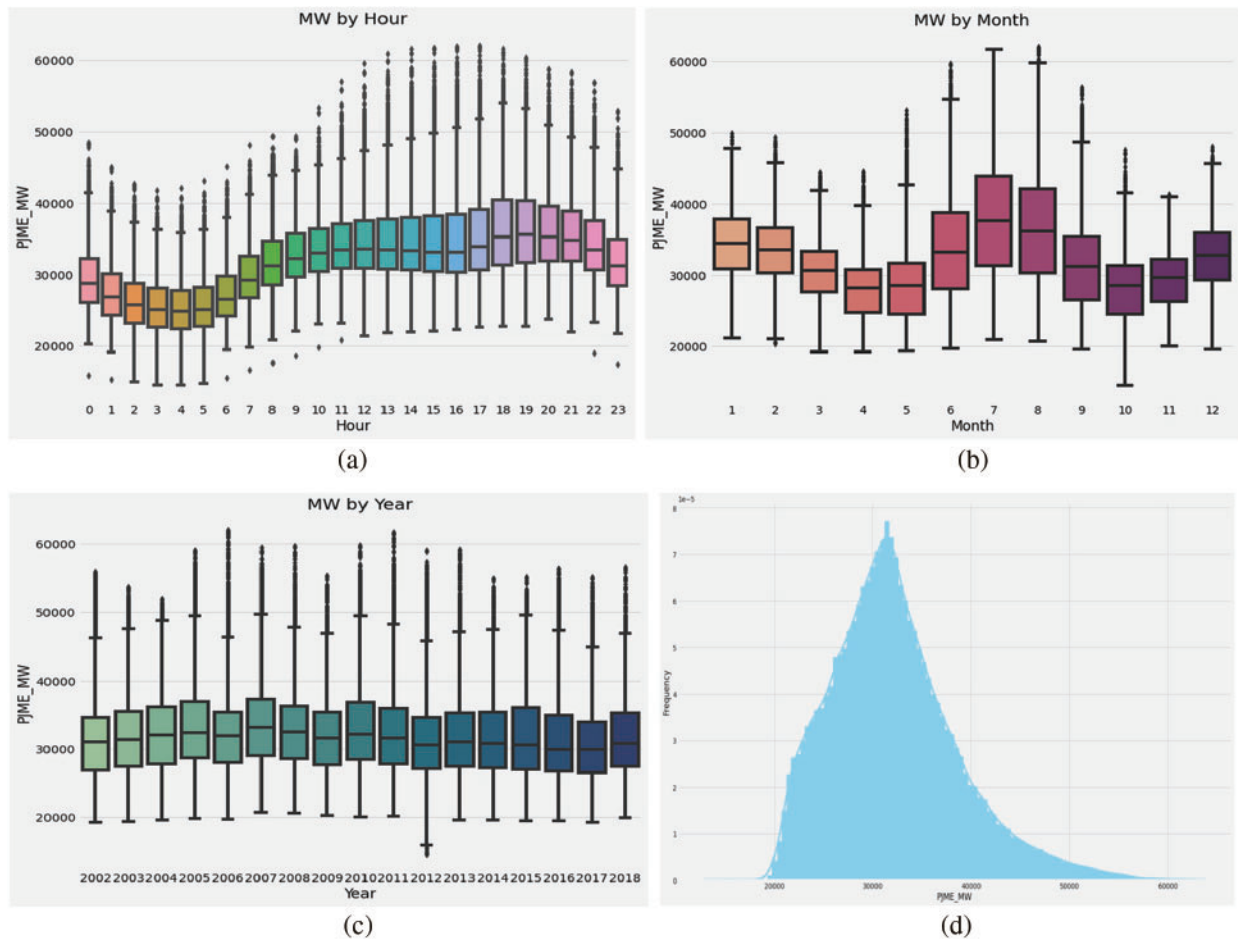
Figure 5: Visualizing the dataset features

Table 1: The metrics used in evaluating the achieved results

Metrics	Formula
MBE	$= \dfrac{1}{N}\sum_{n=1}^{N}\left(\widehat{V}_n - V_n\right)$
MAE	$= \dfrac{1}{N}\sum_{n=1}^{N}\left|\widehat{V}_n - V_n\right|$
NSE	$= 1 - \dfrac{\sum_{n=1}^{N}\left(V_n - \widehat{V}_n\right)^2}{\sum_{n=1}^{N}\left(V_n - \overline{\widehat{V}_n}\right)^2}$
RRMSE	$= \dfrac{RMSE}{\overline{V}_m} \times 100$

(Continued)

Table 1: Continued

Metrics	Formula
RMSE	$= \sqrt{\dfrac{1}{M}\sum_{m=1}^{M}\left[\widehat{V_m} - V_m\right]^2}$
R2	$= 1 - \dfrac{unexpected\ variation}{Total\ variation}$
WI	$= 1 - \dfrac{\sum_{m=1}^{M}\left|\widehat{V_m} - V_m\right|}{\sum_{m=1}^{M}\left|V_m - \overline{V_m}\right| + \left|\widehat{V_m} - \overline{\widehat{V_m}}\right|}$
r	$= \dfrac{\sum_{m=1}^{M}\left(\widehat{V_m} - \overline{\widehat{V_m}}\right)\left(V_m - \overline{V_m}\right)}{\sqrt{\left[\sum_{m=1}^{M}\left(\widehat{V_m} - \overline{\widehat{V_m}}\right)^2\right]\left[\sum_{m=1}^{M}\left(V_m - \overline{V_m}\right)^2\right]}}$

The evaluation of these criteria based on the achieved results using the proposed model and other prediction models is presented in Table 2. As shown in this table, the proposed approach achieves better values than the other prediction models for all types of evaluation criteria, which proves the superiority of the proposed model in predicting the energy consumption accurately.

Table 2: The results of the evaluation criteria using the proposed method and other competing methods

	SVR	KNN	RF	MLP	LSTM	DDTOSFS + LSTM
RMSE	0.049	0.012	0.026	0.010	0.004	0.000
RRMSE	70.56	22.05	40.99	11.19	11.19	6.227
WI	0.605	0.913	0.801	0.901	0.954	0.971
MBE	−0.057	−0.008	−0.016	−0.017	−0.007	−0.007
r	0.993	0.979	0.946	0.992	0.992	0.994
R2	0.991	0.963	0.900	0.989	0.988	0.993
NSE	0.646	0.961	0.877	0.970	0.987	0.993
MAE	0.048	0.007	0.022	0.008	0.001	−0.001

Five other optimization algorithms are incorporated in the experiments to prove the superiority of the proposed optimized LSTM. These algorithms are GWO [23], PSO [24], WOA [25], SFS [26], and DTO [27]. These algorithms are used to optimize the LSTM-based model parameters. The recorded evaluation results are presented in Table 3 in terms of the prediction results statistical analysis. As represented by the results in the table, the proposed approach is shown superior when compared to other optimization algorithms.

Table 3: Analysis of the results achieved by LSTM when optimized using the proposed optimization algorithm and other optimization algorithms

	DDTOSFS + LSTM	DTO + LSTM	SFS + LSTM	PSO + LSTM	GWO + LSTM	WOA + LSTM
Range	0.000011	0.00014	0.0002	0.0001483	0.00011	0.0003
Median	0.0001381	0.0004115	0.000537	0.0006387	0.0007115	0.0008872
75% percentile	0.0001381	0.0004115	0.000537	0.0006387	0.0007115	0.0008901
Std. error of mean	0.000001016	0.00001118	0.00002134	0.00001158	0.000008222	0.00002507
Minimum	0.0001281	0.0003115	0.000337	0.0005387	0.0006612	0.0006872
Std. deviation	0.000003213	0.00003534	0.00006749	0.00003661	0.000026	0.00007929
90% percentile	0.000139	0.0004475	0.000537	0.0006822	0.0007652	0.0009784
Actual confidence level	97.85%	97.85%	97.85%	97.85%	97.85%	97.85%
Upper confidence limit	0.0001381	0.0004115	0.000537	0.0006387	0.0007115	0.0008987
10% percentile	0.0001291	0.0003215	0.000347	0.0005487	0.0006662	0.0006972
Mean	0.0001372	0.0004055	0.000507	0.0006335	0.0007124	0.0008684
25% percentile	0.0001381	0.0004115	0.000512	0.0006387	0.0007115	0.0008622
Maximum	0.0001391	0.0004515	0.000537	0.000687	0.0007712	0.0009872
Number of values	10	10	10	10	10	10
Lower confidence limit	0.0001381	0.0004115	0.000437	0.0006387	0.0007115	0.0007872

Fig. 6 displays four plots to illustrate the model performance from the point of view of a visual representation of the prediction results achieved by the proposed approach. The predicted energy consumption is mapped against the residual error plot shown in Fig. 6a, and the homoscedasticity plot shown in Fig. 6b. The small size of the residual errors in these plots is an indication of the reliability of the projected values. Fig. 6c is a quartile-quartile (QQ) plot comparing the predicted and observed values, illustrating how well they match up [28–35]. To see that the results closely follow a straight line, as shown in the figure, is conclusive evidence for the validity of the suggested model. Prediction errors are visualized using the heatmap as in Fig. 6d. From these plots, the proposed DDTOSFS + LSTM approach is proven to provide the least amount of inaccuracy in forecasting energy consumption.
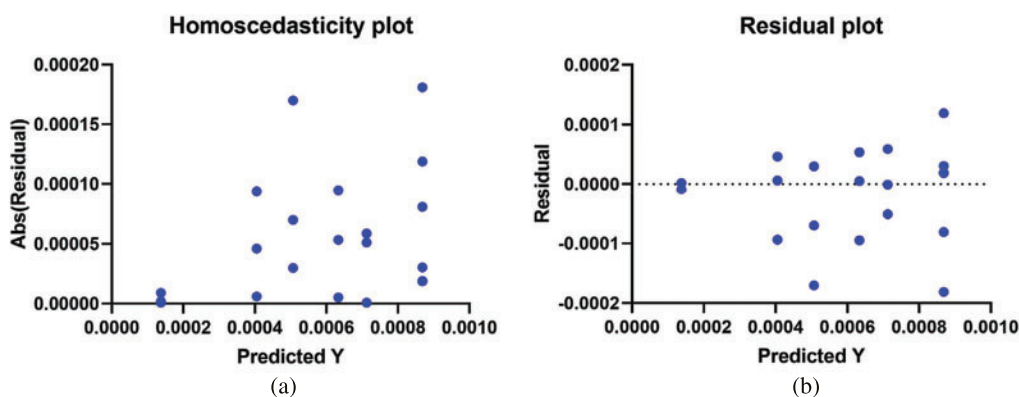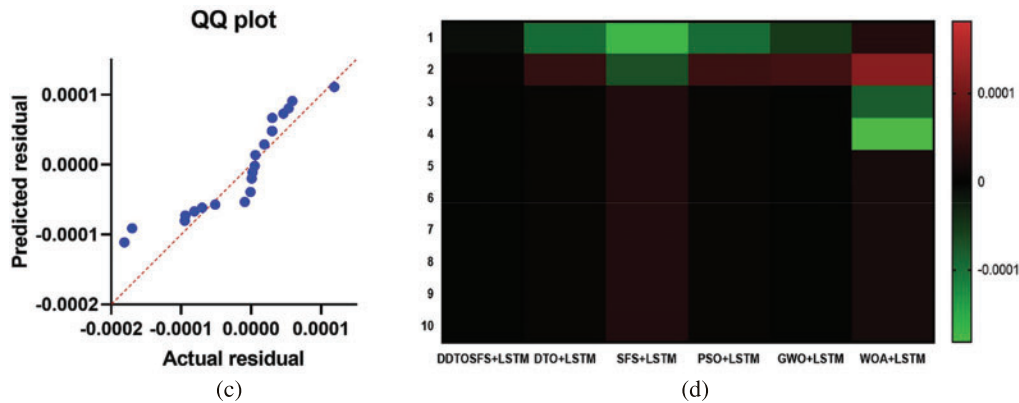

(a)

(b)

Figure 6: (Continued)

Figure 6: Representation of the results achieved using the proposed method

In addition, the suggested DDTOSFS + LSTM approach is compared to the other method using a statistical difference to determine the p-values, demonstrating the substantial difference between the two methods. Wilcoxon's rank-sum test is used to carry out this analysis. Both null and alternative hypotheses are posited before the test is run. Average algorithmic parameters are equalized under the null hypothesis H0 (DDTOSFS + LSTM = DTO + LSTM, SFS + LSTM = PSO + LSTM, GWO + LSTM = WOA + LSTM). Alternatively, H1 proposes that the algorithms' means are not comparable. Table 5 displays the results of the Wilcoxon rank-sum test. The table shows a statistically significant difference between the suggested DDTOSFS + LSTM algorithm and the other algorithms, with p-values of less than 0.05.

The ANOVA test is also used to further investigate the method's efficacy. In this analysis, two basic hypotheses are posed: the alternative and null hypotheses. Mean values of the algorithm are equalized, DDTOSFS + LSTM = DTO + LSTM = SFS + LSTM = PSO + LSTM = GWO + LSTM = WOA + LSTM, for the null hypothesis denoted by H0. Whereas, under the alternative hypothesis H1, the means of the algorithms are not comparable. Table 4 displays the findings of the analysis of variance. The table shows that the suggested algorithm performs as predicted when compared with alternative feature selection techniques. The root-mean-squared error (RMSE) between the proposed technique and the other methods' predictions is displayed in Fig. 7. You can see that the proposed model has the lowest RMSE values in this graph. The prediction error histogram is displayed in Fig. 8. As can be seen in the image, the suggested model yields the lowest error values in its predictions compared to the other techniques. These figures underline the efficacy of the proposed strategy for estimating energy needs. The comparison between actual and expected energy use is also displayed in Fig. 9. To illustrate the reliability of the proposed strategy, the picture above superimposes projected and observed energy consumption.

Table 4: The results of the one-way ANOVA test based on the results achieved by LSTM when optimized using the proposed optimization algorithm and other optimization algorithms

	SS	DF	MS	F (DFn, DFd)	P value
Treatment	3.28E-06	5	6.55E-07	F (5, 54) = 278.4	$P < 0.0001$
Residual	1.27E-07	54	2.35E-09		
Total	3.4E-06	59			

Table 5: Wilcoxon signed-rank test applied to the results of the optimization methods

|	DDTOSFS + LSTM	DTO + LSTM	SFS + LSTM	PSO + LSTM	GWO + LSTM	WOA + LSTM
Theoretical median	0	0	0	0	0	0
Number of values	10	10	10	10	10	10
Actual median	0.0001381	0.0004115	0.000537	0.0006387	0.0007115	0.0008872
Discrepancy	0.0001381	0.0004115	0.000537	0.0006387	0.0007115	0.0008872
Sum of negative ranks	0	0	0	0	0	0
P value (two tailed)	0.002	0.002	0.002	0.002	0.002	0.002
Sum of positive ranks	55	55	55	55	55	55
Significant (alpha = 0.05)?	Yes	Yes	Yes	Yes	Yes	Yes
Sum of signed ranks (W)	55	55	55	55	55	55
P value summary	**	**	**	**	**	**
Exact or estimate?	Exact	Exact	Exact	Exact	Exact	Exact
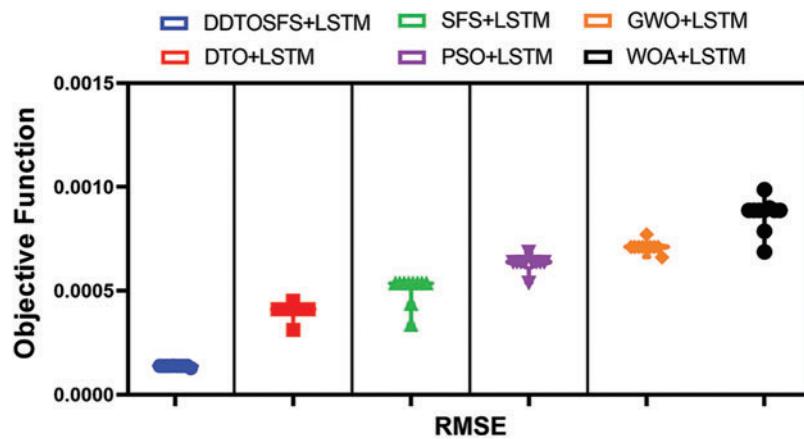
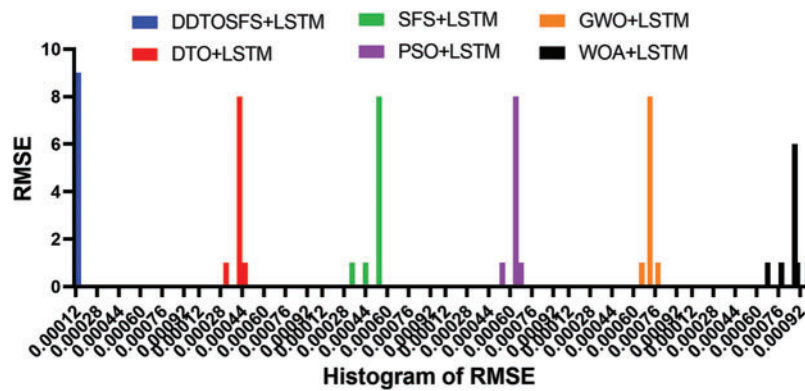**Figure 7:** RMSE values achieved by LSTM when optimized using different algorithms

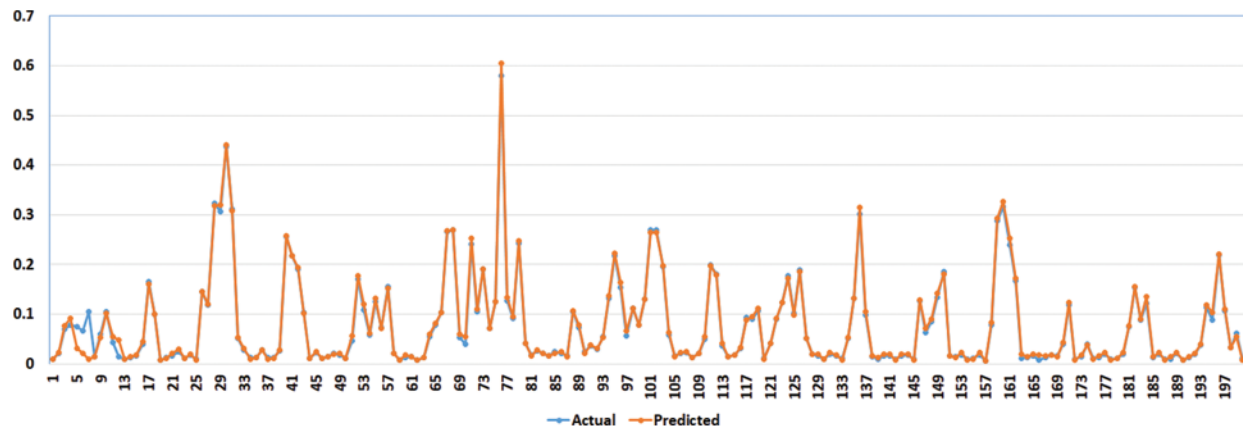**Figure 8:** Histogram of RMSE values achieved by LSTM model optimized using optimization algorithms

Figure 9: The predicted *vs.* actual energy consumptions resulting using the proposed model

5 Conclusion

In this work, we proposed a new strategy for forecasting energy consumption. The proposed approach relies on improving the performance of LSTM by adjusting its parameters, which allows for precise prediction results. Multiple optimization strategies and machine learning models are used to evaluate the proposed technique against a vast group of methods. Statistical analysis is used to demonstrate the reliability and validity of the suggested approach, which is validated by the findings. Additionally, a series of plots, including QQ, residual, homoscedasticity, and histograms, are used to conduct an in-depth study of the final results. This evaluation verified the efficiency and excellence of the suggested approach to precisely forecasting energy usage. The findings of this research can help in decision making and risk management considered by the governing authorities.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] S. Das, T. Choudhury and S. Mohapatra, "Data analytics to increase efficiency of the AI based energy consumption predictor," in *Int. Conf. on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE)*, Keonjhar, India, pp. 1–4, 2020.

[2] S. Jia, W. Yao and Z. Wang, "Accurate prediction method for energy consumption in automatic tool changing process of computer numerical control lathe," in *World Conf. on Mechanical Engineering and Intelligent Manufacturing (WCMEIM)*, Shanghai, China, pp. 241–244, 2019.

[3] G. Sha and Q. Qian, "Prediction of commercial building lighting energy consumption based on EPSO-BP," in *Int. Conf. on Control, Automation and Systems (ICCAS)*, PyeongChang, Korea (South), pp. 1035–1040, 2018.

[4] P. Schirmer, I. Mporas and I. Potamitis, "Evaluation of regression algorithms in residential energy consumption prediction," in *European Conf. on Electrical Engineering and Computer Science (EECS)*, Athens, Greece, pp. 22–25, 2019.

[5] X. Wang, X. Li, D. Qin, Y. Wang, L. Liu *et al.,* "Prediction of industrial power consumption and air pollutant emission in energy internet," in *Asia Energy and Electrical Engineering Symp. (AEEES)*, Chengdu, China, pp. 1155–1159, 2021.

[6] Z. Wu and W. Chu, "Sampling strategy analysis of machine learning models for energy consumption prediction," in *Int. Conf. on Smart Energy Grid Engineering (SEGE)*, Oshawa, Canada, pp. 77–81, 2021.

[7] M. Fayaz and D. Kim, "A prediction methodology of energy consumption based on deep extreme learning machine and comparative analysis in residential buildings," *Electronics*, vol. 7, no. 10, pp. 222, 2018.

[8] R. Akter, J. -M. Lee and D. -S. Kim, "Analysis and prediction of hourly energy consumption based on long short-term memory neural network," in *Int. Conf. on Information Networking (ICOIN)*, Jeju Island, Korea (South), pp. 732–734, 2021.

[9] J. Li, "Evolutionary modeling of factors affecting energy consumption of urban residential buildings and prediction models," in *Int. Conf. on Electronics, Communication and Aerospace Technology (ICECA)*, Coimbatore, India, pp. 1333–1336, 2021.

[10] Z. Qiuhong, Z. Zhan and W. Jiayi, "Application of regional building energy consumption prediction model in building construction," in *Int. Conf. on Intelligent Transportation, Big Data & Smart City (ICITBS)*, Vientiane, Laos, pp. 92–94, 2020.

[11] Y. Liang and Z. Hu, "Prediction method of energy consumption based on multiple energy-related features in data center," in *Int. Conf. on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking*, Xiamen, China, pp. 140–146, 2019.

[12] M. Chuang, W. Yikuai, Z. Junda, C. Ke, G. Feixiang *et al.,* "Research on user electricity consumption behavior and energy consumption modeling in big data environment," in *Int. Conf. on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, Nanchang, China, pp. 220–224, 2021.

[13] X. Zhao, "Energy saving prediction method for public buildings based on data mining," in *Int. Conf. on Measuring Technology and Mechatronics Automation (ICMTMA)*, Beihai, China, pp. 484–487, 2021.

[14] M. Babaoğlu and B. Haznedar, "Turkey long-term energy consumption prediction using whale optimization algorithm," in *Signal Processing and Communications Applications Conf. (SIU)*, Istanbul, Turkey, pp. 1–4, 2021.

[15] F. Janković, L. Šćekić and S. Mujović, "Matlab/Simulink based energy consumption prediction of electric vehicles," in *Int. Symp. on Power Electronics*, Novi Sad, Serbia, pp. 1–5, 2021.

[16] A. Rahman and A. Smith, "Predicting heating demand and sizing a stratified thermal storage tank using deep learning algorithms," *Applied Energy*, vol. 228, no. 1, pp. 108–121, 2018.

[17] I. Goodfellow, Y. Bengio and A. Courville, "Deep Learning," in *MIT Press*, Cambridge, 2016.

[18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computing*, vol. 9, no. 8, pp. 1735–1780, 1997.

[19] A. Takieldeen, E. -S. M. El-kenawy, E. Hadwan and M. Zaki, "Dipper throated optimization algorithm for unconstrained function and feature selection," *CMC-Computers, Materials & Continua*, vol. 72, no. 1, pp. 1465–1481, 2022.

[20] R. Mulla, K. Banachewicz and A. Dahesh "Hourly Energy Consumption," 2018, Accessed: December 20, 2022. [Online]. Available: https://www.kaggle.com/datasets/robikscube/hourly-energy-consumption

[21] E.-S. M. El-Kenawy, S. Mirjalili, S. Ghoneim, M. Eid, M. El-Said *et al.,* "Advanced ensemble model for solar radiation forecasting using sine cosine algorithm and newton's laws," *IEEE Access*, vol. 9, no. 1, pp. 115750–115765, 2021.

[22] A. Abdelhamid and S. Alotaibi, "Robust prediction of the bandwidth of metamaterial antenna using deep learning," *Computers, Materials & Continua*, vol. 72, no. 2, pp. 2305–2321, 2022.

[23] H. R. Hussien, E. -S. M. El-Kenawy and A. El-Desouky, "EEG channel selection using a modified grey wolf optimizer," *European Journal of Electrical Engineering and Computer Science*, vol. 5, no. 1, pp. 17–24, 2021.

[24] E. -S. M. El-kenawy and M. Eid, "Hybrid gray wolf and particle swarm optimization for feature selection," *International Journal of Innovative Computing, Information & Control*, vol. 16, no. 1, pp. 831–844, 2020.

[25] M. M. Eid, E. -S. M. El-kenawy and A. Ibrahim, "A binary sine cosine-modified whale optimization algorithm for feature selection," in *National Computing Colleges Conf.*, Taif, Saudi Arabia, pp. 1–6, 2021.

[26] N. Gorgolis, I. Hatzilygeroudis, Z. Istenes and L. Gyenne, "Hyperparameter optimization of LSTM network models through genetic algorithm," in *Int. Conf. on Information, Intelligence, Systems and Applications (IISA)*, Patras, Greece, pp. 1–4, 2019.

[27] T. Tran, K. Teuong and D. Vo, "Stochastic fractal search algorithm for reconfiguration of distribution networks with distributed generations." *Ain Shams Engineering Journal*, vol. 11, no. 2, pp. 389–407, 2020.

[28] E. M. El-Kenawy, M. M. Eid, M. Saber and A. Ibrahim, "MbGWO-SFS: Modified binary grey wolf optimizer based on stochastic fractal search for feature selection," *IEEE Access*, vol. 8, pp. 107635–107649, 2020.

[29] N. Abdel Samee, E. M. El-Kenawy, G. Atteia, M. M. Jamjoom, A. Ibrahim *et al.,* "Metaheuristic optimization through deep learning classification of covid-19 in chest x-ray images," *Computers, Materials & Continua*, vol. 73, no. 2, pp. 4193–4210, 2022.

[30] E. -S. M. El-Kenawy, S. Mirjalili, A. Ibrahim, M. Alrahmawy, M. El-Said *et al.,* "Advanced meta-heuristics, convolutional neural networks, and feature selectors for efficient COVID-19 X-ray chest image classification," *IEEE Access*, vol. 9, no. 1, pp. 36019–36037, 2021.

[31] A. Abdelhamid and S. Alotaibi, "Optimized two-level ensemble model for predicting the parameters of metamaterial antenna," *Computers, Materials & Continua*, vol. 73, no. 1, pp. 917–933, 2022.

[32] D. Sami Khafaga, A. Ali Alhussan, E. M. El-kenawy, A. E. Takieldeen, T. M. Hassan *et al.,* "Meta-heuristics for feature selection and classification in diagnostic breast cancer," *Computers, Materials & Continua*, vol. 73, no. 1, pp. 749–765, 2022.

[33] D. Sami Khafaga, A. Ali Alhussan, E. M. El-kenawy, A. Ibrahim, S. H. Abd Elkhalik *et al.,* "Improved prediction of metamaterial antenna bandwidth using adaptive optimization of LSTM," *Computers, Materials & Continua*, vol. 73, no. 1, pp. 865–881, 2022.

[34] E. -S. M. El-Kenawy, S. Mirjalili, F. Alassery, Y. Zhang, M. Eid *et al.,* "Novel meta-heuristic algorithm for feature selection, unconstrained functions and engineering problems," *IEEE Access*, vol. 10, pp. 40536–40555, 2022.

[35] A. Abdelhamid, E. -S. M. El-kenawy, B. Alotaibi, M. Abdelkader, A. Ibrahim *et al.,* "Robust speech emotion recognition using CNN + LSTM based on stochastic fractal search optimization algorithm," *IEEE Access*, vol. 10, pp. 49265–49284, 2022.