VOL. 23, NO. 3, 2024, 61-66
elektrika.utm.my
ISSN 0128-4428

ELEKTRIKA
Journal of Electrical Engineering

Development of an Optimized Neural Network Model using Hyperparameters Optimization for Electrical Load Prediction

Abdulsemiu A. Olawuyi[1*], **Titus O. Ajewole**[1], **Olatunde Oladepo**[1] and **Tolulope Awofolaju**[1]

[1]Department of Electrical and Electronic Engineering, Osun State University, Osogbo, Nigeria

*Corresponding author: abdulsemiu.olawuyi@uniosun.edu.ng

Abstract: Electrical load prediction has become essential to the efficient operation, control and management of modern electric power systems. Various machine learning prediction models have been developed for electrical load prediction, this includes Support Vector Regression, Fuzzy Logic and Neural Network (NN) modelling approaches. However, incorrect selection of model hyperparameters, which are parameters that affect the output of the prediction models, could result in low prediction accuracy of machine learning models. Hence, the development of an optimized NN model for electrical load prediction was presented in this study. Historical daily data of temperature, rainfall, relative humidity and windspeed for Osogbo, Nigeria was obtained from the National Aeronautics and Space Administration website; while electrical load data for the same location was collected from the Transmission Company of Nigeria. The data captured a period of five years (2017 to 2021). The NN models were developed with MATLAB R2022a software, and two hyperparameters, hidden layers and neuron counts, were optimized using the Bayesian optimization technique to enhance the quality of the models. The models were evaluated using mean absolute error (MAE), and root mean square error (RMSE). The MAE and RMSE for the non-optimized NN model were 6.5247 and 8.2725 respectively. Meanwhile, for the optimized NN model, the MAE and RMSE were 5.6571 and 7.4289 respectively. The obtained results show that the optimized NN performed better than the non-optimized NN models. Therefore, for more accurate load prediction, the method developed of this research is suggested for use by utility providers.

Keywords: Bayesian optimization, hyperparameters, load prediction, machine learning, neural network

Article History: received 28 June 2024; accepted 10 November 2024; published 30 December 2024

1. INTRODUCTION

The importance of electrical power in the development of any society cannot be overemphasized. The demand for electrical has increased owing to exponential increase in human population as well as advancement in science and technology [1]. Resources have been invested by the stakeholders in the power industry to sustain the quality and control of electrical power. Modern power system needs to be stable in order to get the desired results of supply to the markets. Energy prediction is essential in maintaining and controlling modern power grid system especially with the addition of green energy into the grid. Researchers employing several methods and approaches, have carried-out load predictions from different time horizons. Short-term load prediction is the least time horizons. The prediction is from minutes to weeks. Short-term load forecast is essential in electrical system operation for policy making in energy sector, spinning reserve planning and grid stability. Secondly, the midterm load prediction usually studies the prediction from week to months. Midterm load prediction is employed in power system expansion and equipment procurement.

Machine learning (ML) can be applied in a variety of fields, including healthcare [2], finance [3, 4] education [5] and transportation [6]. ML plays significant roles in power system by improving efficiency, reliability and sustainability [7]. Power system operation can be improved with the application of machine learning and the transmission grid by predicting demand and adjusting power output accordingly. This will help minimize the cost of generating electrical energy and enhance power system overall performance [8]. The addition of the renewable resources like as wind and solar, into the power system can be managed with machine learning. It will minimize the use of fossil fuel and improve and further incorporation of eco-friendly energy. Also, future energy demand could be known with respect to past dataset with the aid of Machine learning. This can assist the policy makers in the energy sector to plan ahead and also make informed decision relating to how to minimize waste and improve the overall performance indices of the energy sector [9]. There are challenges related with the use of ML in power system. Machine learning algorithms rely on large amount of data for training and operation [10]. Dataset availability in the power sector could be challenging and when it is available, there may be problem of integrity due to missing values and noise. In addition, current power system control and monitoring systems have a hard time integrating AI models. It could take a lot of work and complexity to implement the integration [11]. Failures in the power system can lead to significant effects such as power outages and equipment damage. However, ensuring the reliability and safety of the models can be problematic due to the uncertain and dynamic nature of the power system design. Load prediction in power systems refers to the process of forecasting the future demand for electricity [12].

Accurate load prediction is important for utilities and power system operators, as it helps them to deal with the problem of insufficient availability of electrical energy and infrastructure to meet the forecasted energy required [13]. There are several approaches to load prediction, including statistical, data-driven and physic-based models [14, 15] . Statistical models use historical data to forecast future loads based on past trends and patterns [16, 17]. Data-driven models like artificial neural networks, use machine learning techniques to make predictions based on large dataset of input variables [18]. Physics-based models use physical principles and equations to simulate the behavior of the power system and forecast future loads [19]. Load prediction methods include time series [20], neural networks [21], support vector regression [22], random forests [23], decision trees [24] regression analysis [25] and adaptive neuro-fuzzy inference system [26].

Three neural network algorithms were investigated by [27] . The algorithms were developed using MATLAB. The three algorithms were Levenberg-Marquardt (LM), gradient descent with momentum and gradient descent algorithm. The neural network had fifteen neurons with a hidden layer, six inputs and an output. The best performing algorithm in their work was LM algorithm. The LM algorithm had 57.9962 and 6.4675 as Mean Square Error (MSE) and Mean Absolute Error (MAE) respectively. [28] proposed novel data pre-processing strategies for training a Neural Network (NN) for electrical energy prediction. Two essential preprocessing methods were recommended, focusing on the importance of specific input parameters with respect to its output values, resulting in improved prediction results when compared to the classic methods.

The paper emphasizes the critical role of hyperparameters in determining the performance of machine learning models, particularly for neural networks. A fair evaluation of any machine learning model requires careful investigation and tuning of its hyperparameters. In this study, we demonstrate that by optimizing the hyperparameters of a neural network model, where the accuracy of electrical load forecasting algorithms is significantly enhanced, leading to more reliable predictions.

This study is organized thusly: methods and formulation of neural network model employed were describes in Section II, the results and discussion are presented in following Section and Section IV concludes the research presented.

2. DATA AND METHODOLOGY

2.1 Data Acquisition

The electrical dataset employed in this study were obtained from the Transmission Company of Nigeria (TCN) and the weather parameters was collected from National Aeronautics and Space Administration (NASA) website. The electrical load dataset was daily dataset from 2017 to 2021. Osogbo, an ancient city in Osun State, is the capital city. The city is situated on $7^{\circ}46^{'}N$ longitude and $4^{\circ}34^{'}E$ latitude (Taiwo, *et al*., 2019). The population of the state is 4,705,589 (National Population Commission, 2006) and land mass area of 9,026 sq. km. (Funke, 2008).

2.2 Data Preprocessing

One of the important stages in machine learning is the preprocessing of data, as it ensures that the data is formatted in a way that is suitable for optimal analysis and modeling. This study utilizes data cleansing, data formatting, and data exploration techniques. Data formatting involves altering the structure of data to ensure its compatibility with analysis or modeling purposes. The data formatting method utilized in this research is normalization. The data in this study was normalized using the min-max scaling technique.

2.3 Bayesian Optimization

The Bayesian optimization technique was applied for optimizing the hyperparameters of the Neural Network (NN) model. Bayesian optimization is a highly effective method for optimizing machine learning models. Bayesian optimization efficiently manages the trade-off between exploration and exploitation. The process entails constructing a proxy model of the objective function. This model facilitates the identification of favorable regions within the search space and directs the search towards locations that are more likely to produce excellent solutions in an efficient manner. Bayesian optimization is applied in various domains, ranging from machine learning where it is used for hyperparameter tuning, to finance where it is used for resource allocation. Its capacity to effectively manage goal functions that are both noisy and costly makes it well-suited for real-world optimization difficulties.

2.4 Neural Network

Neural Network (NN) architecture was developed using MATLAB R2022a. NN has various connected neurons that aid information transmission. 70% and 30% of the dataset was employed for training and testing respectively. The input data was fed through the NN, and the models was compared to the actual output. Based on the difference between predicted and actual output, the NN model adjusts its internal biases and weights to improve its performance. Figure 1 shows the custom view of NN in MATLAB environment.

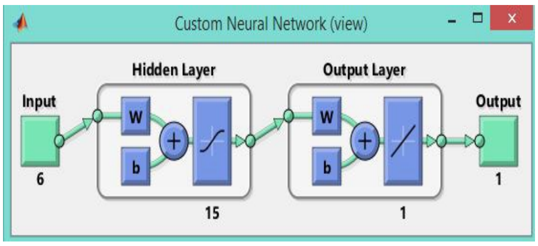

Figure 1. Neural network

2.4.1 Neural Network Mathematical Model

The mathematical model of NN was expressed using Equations (1) to (4), from the input section to the output section of the NN network. For an NN model with an inner layer and an output layer, input to the inner layer is expressed with Equations (1) and (2).

$$v^{(1)} = W^{(1)}x + b^{(1)} \tag{1}$$

$$d^{(1)} = \Phi^{(1)}(v^{(1)}) \tag{2}$$

Equations (3) and (4) represent the NN hidden layer

$$v^{(2)} = W^{(2)}d^{(1)} + b^{(2)} \tag{3}$$

$$\hat{y} = \Phi^{(2)}(v^{(2)}) \tag{4}$$

x is the input vector to the neural network, $b^{(l)}$ is the bias l vector of layer, vector of inputs, as well as weights plus biases, to the neurons in the layer l is $v^{(l)}$, $\Phi^{(l)}$ is the activation function, Rectified Linear Unit (ReLU) used in layer l. $\hat{y}$ represents the output of the NN model. The ReLU function is shown in Figure 2.

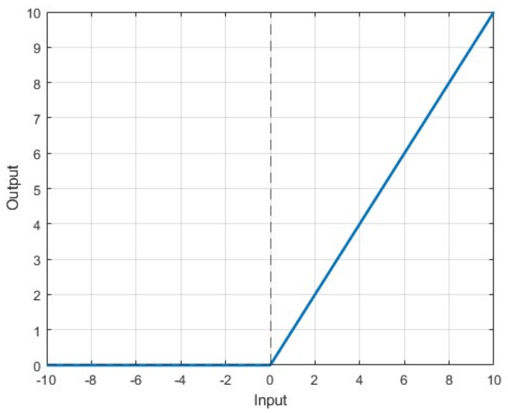

Figure 2. ReLU activation function

2.4.2 Optimized Neural Network (ONN) Mathematical Model

There is need to obtain the optimal hyperparameters of an NN model that optimizes the prediction error on the test set while satisfying the constraints. The constraints optimization was formulated as indicated in Equation (5).

$$\left.\begin{array}{l} \quad Min_{\gamma}\ MAE(\gamma) \\ \quad Min_{\gamma}\ RMSE(\gamma) \\ subject\ to \\ \quad 1 \le number\ of\ layer \le 5 \\ \quad 10 \le neuron\ counts \le 100 \\ where \\ \quad number\ of\ layer,\ neuron\ counts > 0 \end{array}\right\} \tag{5}$$

γ is the hyperparameters of the NN model. Equations (6) and (7) represent the MAE and RMSE respectively.

$$MAE(\gamma) = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \left| y_i - f_{\gamma}(x_i) \right| \tag{6}$$

$$RMSE(\gamma) = \sqrt{\frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \left(y_i - f_{\gamma}(x_i) \right)^2} \tag{7}$$

$f_{\gamma}(x_i)$ is the prediction of the NN model with hyperparameters, γ, for input x_i.

Figure 3 shows the optimized neural network model. After loading the dataset in the MATLAB R2022a environment, the objective function with respect to the hyperparameters were defined. Bayesian optimization was run to optimized the hyperparameters and also minimizing the performance metrics employed as the objective function of the model. 100 iterations were performed and the ten best hyperparameters were recorded. The best combination of the hyperparameters were selected and evaluated using MAE and RMSE.

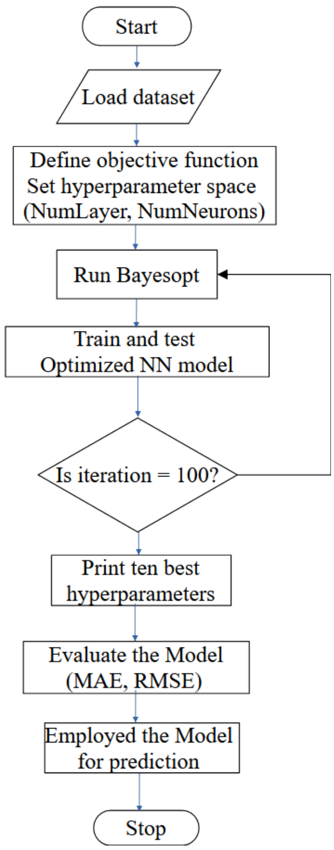

Figure 3. ONN flowchart

3. RESULT AND DISCUSSION

The neural network models were developed based on the models' parameters describes in Section II. Both the non-optimized and optimized neural network were compared with respect to the performance metrics.

3.1 Non-optimized neural network

Figure 4 represents the prediction result of the developed neural network model. The model employed the default hyperparameters of 10 neurons and two hidden layers with

ReLU as activation function. The MAE and RMSE for the non-optimized NN model were 6.5247 and 8.2725 respectively. The results indicates that the prediction model performed fairly. However, the results could further be enhanced with hyperparameter optimization.

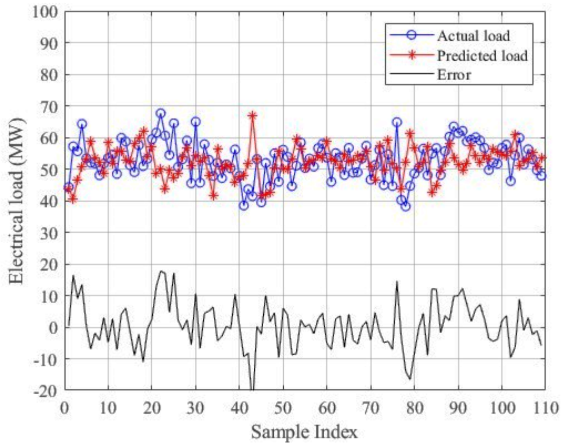

Figure 4. Prediction using non-optimized NN model

3.2 Optimized neural network model

Employing the optimization function describe in Equation (5) and the minimizing the error function with Equations (6) and (7). Figure 5 depicts the optimized neural network. The optimized model shows an improvement of the prediction of the electrical load.

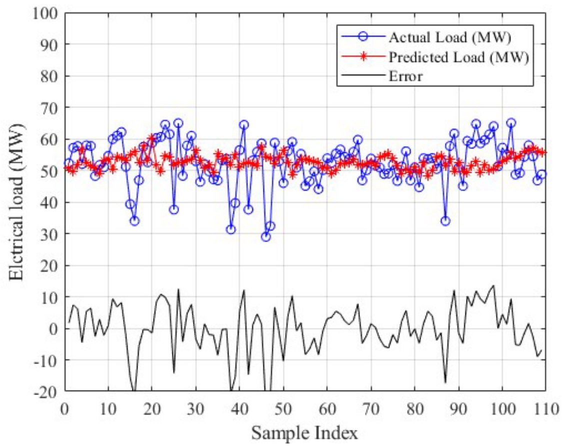

Figure 5. Prediction using optimized NN model

The MAE and RMSE for the optimized NN model were 5.6571 and 7.4289 respectively.

Table 1 shows the best 10 performances of the optimized NN model with MAE as performance metric. The most optimized has number of layer and neurons count of 3 and 67 respectively. Figure 6 depicts the optimized NN objective function model with optimized MAE value of 5.6571.

Table 2 illustrates the best 10 performances of the optimized NN model with RMSE as performance metric. The most optimized has number of layer and neurons count of 2 and 82 respectively. Figure 7 shows the optimized NN objective function model with optimized RMSE value of 7.4289.

Figures 6 and 7 show the objective function of the NN model using MAE and RMSE respectively. Two objective variables, neuron counts and number of layers are plotted against the estimated objective function value. The objective function is MAE and RMSE in Figures 6 and 7 respectively. The total number of iterations are 100, represented as observed points in Figures 6 and 7. Each of the iteration point was evaluated. The iteration point with the least estimated objective function value was selected and subsequently evaluated with MAE and RMSE

Table 1. Top 10 performance metrics and hyperparameters using MAE

numLayer	numNeurons	MAE
3	67	5.6571
3	63	5.6949
1	57	5.7011
5	76	5.7079
4	57	5.7123
3	87	5.7157
2	100	5.72
2	100	5.7229
2	73	5.7241
3	15	5.7256

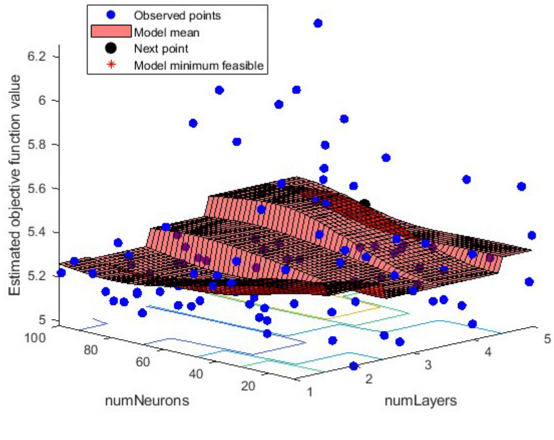

Figure 6. Optimized NN objective function model (MAE)

Figure 7 shows the performance metrics comparison between non-optimized NN and optimized NN model. Figure 7 further highlight the crucial role that the hyperparameters optimization play in the overall improvement of the machine learning model.

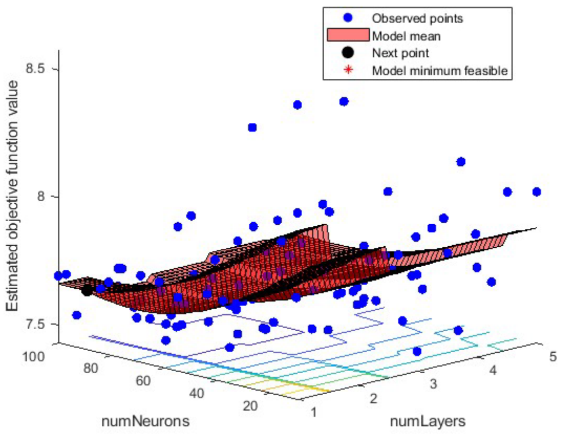

Figure 7. Optimized NN objective function model (RMSE)

Table 2. Top 10 performance metrics and hyperparameters using RMSE

numLayer	numNeurons	RMSE
2	82	7.4289
5	75	7.4465
3	68	7.4572
2	58	7.4595
3	64	7.4892
2	93	7.4913
2	78	7.4919
5	95	7.4925
2	81	7.4935
2	88	7.4998

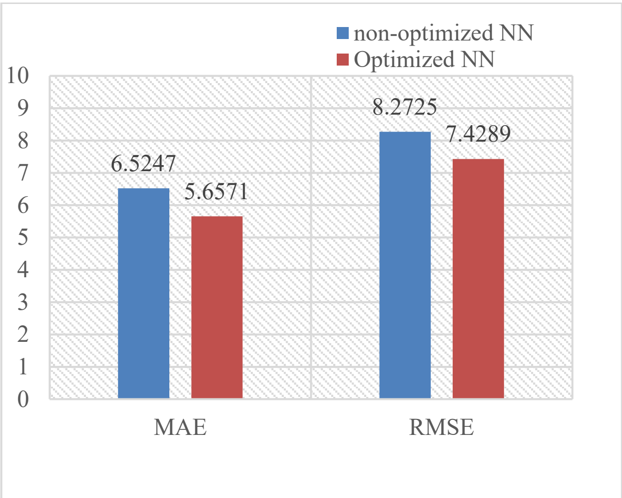

Figure 8. Performance metrics comparison between non-optimized NN and optimized NN model

Figure 8 shows that a better performance of machine model could achieved as a result optimal hyperparameters optimization. This is evident in more evident in Figure 8. The impact of the finding revealed that, to achieved a better electrical loaf prediction, hyperparameters need to be at optimal value.

NN models do not perform well when the datasets are small. When larger are unavailable, support vector machine could be a better candidate for model development. The neural network implemented in this study could further be improved by incorporating other NN hyperparameters like activation function and epoch number.

4. CONCLUSION

In summary, this paper has highlighted the importance of hyperparameters tuning in improving the performance of machine learning models, specifically neural network. The findings demonstrate that performance of any machine learning models cannot be fully capture without investigating and tuning its hyperparameters. Other NN hyperparameters like activation functions and number of epochs could be investigated in further studies. Future studies can also consider the effects of hyperparameters of other machine learning models like support vector regression and k-nearest neighbor towards improved electrical load prediction.

REFERENCES

[1] A. Latif, S. S. Hussain, D. C. Das, and T. S. Ustun, "State-of-the-art of controllers and soft computing techniques for regulated load frequency management of single/multi-area traditional and renewable energy based power systems," *Appl. Energy*, vol. 266, p. 114858, 2020.

[2] S. Kaushik, "AI in healthcare: time-series forecasting using statistical, neural, and ensemble architectures. Front. Big Data 3 (2020)." 2020.

[3] N. R. Tadapaneni, "Artificial intelligence in finance and investments," *Int. J. Innov. Res. Sci. Eng. Technol.*, vol. 9, no. 5, 2019.

[4] L. Cao, "AI in Finance: Challenges, Techniques, and Opportunities," *ACM Comput. Surv. CSUR*, vol. 55, no. 3, pp. 1–38, 2022.

[5] F. Ouyang and P. Jiao, "Artificial intelligence in education: The three paradigms," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100020, 2021.

[6] L. S. Iyer, "AI enabled applications towards intelligent transportation," *Transp. Eng.*, vol. 5, p. 100083, 2021.

[7] A. K. David and Zhao Rongda, "Advances in AI applications for power system planning." 2019.

[8] S. Vollert, M. Atzmueller, and A. Theissler, "Interpretable Machine Learning: A brief survey from the predictive maintenance perspective," in *2021 26th IEEE international conference on emerging technologies and factory automation (ETFA)*, IEEE, 2021, pp. 01–08.

[9] A. N. Abdalla *et al.*, "Integration of energy storage system and renewable energy sources based on artificial intelligence: An overview," *J. Energy Storage*, vol. 40, p. 102811, 2021.

[10] Y. Zhao, T. Li, X. Zhang, and C. Zhang, "Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future," *Renew. Sustain. Energy Rev.*, vol. 109, pp. 85–101, 2019.

[11] T. Ahmad *et al.*, “Artificial intelligence in sustainable energy industry: Status Quo, challenges and opportunities,” *J. Clean. Prod.*, vol. 289, p. 125834, 2021.
[12] A. Azeem, I. Ismail, S. M. Jameel, F. Romlie, K. U. Danyaro, and S. Shukla, “Deterioration of electrical load forecasting models in a smart grid environment,” *Sensors*, vol. 22, no. 12, p. 4363, 2022.
[13] I. K. Nti, M. Teimeh, O. Nyarko-Boateng, and A. F. Adekoya, “Electricity load forecasting: A systematic review,” *J. Electr. Syst. Inf. Technol.*, vol. 7, no. 1, pp. 1–19, 2020.
[14] P. Li, Y. Zhang, Y. Zhang, and K. Zhang, “Prediction of electric bus energy consumption with stochastic speed profile generation modelling and data driven method based on real-world big data,” *Appl. Energy*, vol. 298, p. 117204, 2021.
[15] S. Nasiri and M. R. Khosravani, “Applications of data-driven approaches in prediction of fatigue and fracture,” *Mater. Today Commun.*, vol. 33, p. 104437, 2022.
[16] M. A. Hammad, B. Jereb, B. Rosi, and D. Dragan, “Methods and models for electric load forecasting: a comprehensive review,” *Logist. Supply Chain Sustain. Glob. Chall.*, vol. 11, no. 1, pp. 51–76, 2020.
[17] C. Si, S. Xu, C. Wan, D. Chen, W. Cui, and J. Zhao, “Electric load clustering in smart grid: Methodologies, applications, and future trends,” *J. Mod. Power Syst. Clean Energy*, vol. 9, no. 2, pp. 237–252, 2021.
[18] J. Seo, S. Kim, S. Lee, H. Jeong, T. Kim, and J. Kim, “Data-driven approach to predicting the energy performance of residential buildings using minimal input data,” *Build. Environ.*, vol. 214, p. 108911, 2022.
[19] Q. Lu and W. Zhang, “Integrating dynamic Bayesian network and physics-based modeling for risk analysis of a time-dependent power distribution system during hurricanes,” *Reliab. Eng. Syst. Saf.*, vol. 220, p. 108290, 2022.
[20] J. Lee and Y. Cho, “National-scale electricity peak load forecasting: Traditional, machine learning, or hybrid model?,” *Energy*, vol. 239, p. 122366, 2022.
[21] R. Wazirali, E. Yaghoubi, M. S. S. Abujazar, R. Ahmad, and A. H. Vakili, “State-of-the-art review on energy and load forecasting in microgrids using artificial neural networks, machine learning, and deep learning techniques,” *Electr. Power Syst. Res.*, vol. 225, p. 109792, 2023.
[22] J. Li, Y. Lei, and S. Yang, “Mid-long term load forecasting model based on support vector machine optimized by improved sparrow search algorithm,” *Energy Rep.*, vol. 8, pp. 491–497, 2022.
[23] S. Dang, L. Peng, J. Zhao, J. Li, and Z. Kong, “A quantile regression random forest-based short-term load probabilistic forecasting method,” *Energies*, vol. 15, no. 2, p. 663, 2022.
[24] R. Chaganti *et al.*, “Building heating and cooling load prediction using ensemble machine learning model,” *Sensors*, vol. 22, no. 19, p. 7692, 2022.
[25] O. Olabode, O. A. Amole, T. Ajewole, and I. Okakwu, “Medium-term load forecasting in A nigerian electricity distribution region using regression analysis techniques,” *2020 Int. Conf. Math. Comput. Eng. Comput. Sci. ICMCECS*, vol. null, pp. 1–5, 2020, doi: 10.1109/ICMCECS47690.2020.240907.
[26] C. Ghenai, O. A. A. Al-Mufti, O. A. M. Al-Isawi, L. H. L. Amirah, and A. Merabet, “Short-term building electrical load forecasting using adaptive neuro-fuzzy inference system (ANFIS),” *J. Build. Eng.*, vol. 52, p. 104323, 2022.
[27] T. O. Ajewole, A. A. Olawuyi, M. K. Agboola, and O. Onarinde, “A Comparative Study on the Performances of Power Systems Load Forecasting Algorithms,” 2021.
[28] A. I. Arvanitidis, D. Bargiotas, A. Daskalopulu, V. M. Laitsos, and L. H. Tsoukalas, “Enhanced short-term load forecasting using artificial neural networks,” *Energies*, vol. 14, no. 22, p. 7788, 2021.