RESEARCH AND APPLICATION OF THE STACKING ENSEMBLE MODEL IN SHORT-TERM LOAD FORECASTING

Nguyen Anh Tuan[1*], Ngo Minh Duc[2], Ngo Thuy Ngan[3], Tran Thanh Ngoc[1]

[1]Industrial University of Ho Chi Minh City, [2]TNU - University of Technology, [3]Ha Noi Metropolitan University

ARTICLE INFO

Received: 14/01/2025

Revised: 06/3/2025

Published: 07/3/2025

KEYWORDS

Stacking Model
XGBoost
LightGBM
Random Forest
Support Vector Regression

ABSTRACT

This study presents the development and implementation of a Stacking Ensemble model for short-term load forecasting, a critical power system management and operations task. The Stacking Ensemble model integrated multiple machine learning algorithms, including XGBoost, LightGBM, Ridge Regression, Random Forest, and Support Vector Regression, to enhance forecasting accuracy and reduce standard error metrics such as Mean Absolute Error, Mean Absolute Percentage Error, Mean Squared Error, and Root Mean Squared Error. The findings demonstrate that the Stacking Ensemble model outperforms individual models when handling complex and volatile datasets. A comprehensive comparison between the Stacking Ensemble and individual models such as XGBoost, LightGBM, Random Forest, and Support Vector Regression reveals that Stacking improves forecasting accuracy and offers excellent stability across different scenarios. This research validates the Stacking Ensemble as a robust and effective approach for optimizing short-term load forecasting, contributing to enhanced energy management and more efficient power system operations.

NGHIÊN CỨU VÀ ỨNG DỤNG MÔ HÌNH STACKING ENSEMBLE TRONG DỰ BÁO PHỤ TẢI NGẮN HẠN

Nguyễn Anh Tuấn[1*], Ngô Minh Đức[2], Ngô Thúy Ngân[3], Trần Thanh Ngọc[1]

[1]Trường Đại học Công nghiệp Thành phố Hồ Chí Minh, [2]Trường Đại học Kỹ thuật Công nghiệp – ĐH Thái Nguyên, [3]Trường Đại học Thủ đô Hà Nội

THÔNG TIN BÀI BÁO

Ngày nhận bài: 14/01/2025

Ngày hoàn thiện: 06/3/2025

Ngày đăng: 07/3/2025

TỪ KHÓA

Mô hình Stacking Ensemble
XGBoost
LightGBM
Random Forest
Support Vector Regression

TÓM TẮT

Nghiên cứu này trình bày quá trình phát triển và triển khai mô hình Stacking Ensemble để dự báo tải ngắn hạn, một nhiệm vụ quan trọng trong quản lý và vận hành hệ thống điện. Mô hình Stacking Ensemble tích hợp nhiều thuật toán học máy, bao gồm XGBoost, LightGBM, Ridge Regression, Random Forest và Support Vector Regression, để tăng cường độ chính xác của dự báo và giảm các số liệu lỗi phổ biến như Mean Absolute Error, Mean Absolute Percentage Error, Mean Squared Error và Root Mean Squared Error. Các phát hiện chứng minh rằng mô hình Stacking Ensemble hoạt động tốt hơn các mô hình riêng lẻ khi xử lý các tập dữ liệu phức tạp và dễ biến động. So sánh toàn diện giữa Stacking Ensemble và các mô hình riêng lẻ như XGBoost, LightGBM, Random Forest và Support Vector Regression cho thấy Stacking không chỉ cải thiện độ chính xác của dự báo mà còn mang lại tính ổn định cao hơn trong các tình huống khác nhau. Nghiên cứu này xác nhận Stacking Ensemble là một phương pháp mạnh mẽ và hiệu quả để tối ưu hóa dự báo tải ngắn hạn, do đó góp phần nâng cao quản lý năng lượng và vận hành hệ thống điện hiệu quả hơn.

DOI: https://doi.org/10.34238/tnu-jst.11886

* Corresponding author. *Email: nguyenanhtuan@iuh.edu.vn*

1. Introduction

Electricity load forecasting plays a crucial role in the management and operation of power systems, particularly as electricity demand continues to increase and fluctuate due to economic development, industrialization, and urbanization. Accurate load forecasting not only optimizes the operation of power systems, ensuring a stable power supply, but also reduces operational costs and maximizes the efficient utilization of energy resources. Over the years, numerous load forecasting models have been developed, including both traditional and advanced machine learning methods. Conventional approaches like Linear Regression [1], [2] and AutoRegressive Integrated Moving Average (ARIMA) [3], [4] rely on historical data for predictions. Meanwhile, modern machine learning models such as Random Forest (RF) [5], [6] and Support Vector Regression (SVR) [7], [8] can handle complex data patterns. Boosting techniques like XGBoost [9], [10] and LightGBM [11], [12] are widely used due to their high performance. Additionally, neural networks such as Artificial Neural Networks (ANN) [13], [14], Convolutional Neural Networks (CNN) [15], [16], Recurrent Neural Networks (RNN) [17], [18], and especially Long Short-Term Memory (LSTM) [19], [20], have shown strong results in time-series forecasting. Other models, including Decision Tree Regression (DTR) [21], [22], Gradient Boosting Machine (GBM) [23], [24], and CatBoost [24], [25], continue to be improved for more accurate load forecasting.

Traditional models are generally easier to understand and implement but may struggle with large and complex datasets. In contrast, modern machine learning models excel at handling complex and non-linear data patterns but require significant computational resources and face challenges in interpretability. Models such as ANN [26], RNN [27], LSTM, and CNN [28] have proven effective in capturing non-linear and sequential patterns in electricity load time-series data. ANN is commonly used to model non-linear relationships between inputs and outputs. RNN, with its recurrent structure, handles sequential data well by retaining information from previous time steps. LSTM, an improvement over RNN, addresses the vanishing gradient problem and is better suited for more extended time-series data. Although initially designed for image processing, CNN has been successfully applied to time-series forecasting due to its ability to extract local features.

However, these models have notable limitations. Hyperparameter tuning is complex, requiring significant time and expertise. Furthermore, they are prone to overfitting when training data is small or noisy, resulting in poor generalization to new data. Complex models like LSTM and CNN also demand substantial computational resources, challenging their practical implementation.

The Stacking Ensemble model was introduced in [29] and [30] to address this challenge. Stacking is a method that integrates multiple forecasting models to construct a composite model, thereby enhancing forecasting accuracy and generalization. Unlike traditional standalone models, stacking exploits the advantages of various base models while mitigating their limitations. This approach leads to improved performance metrics and greater adaptability to dynamic conditions, effectively addressing common challenges associated with ANN, RNN, LSTM, and CNN, particularly regarding computational efficiency and stability.

This study utilizes models such as XGBoost, LightGBM, Random Forest, and SVR to construct a Stacking Ensemble model to enhance the efficiency of electricity load forecasting. This approach reduces common errors such as MAE, MAPE, MSE, and RMSE and provides excellent stability under varying conditions. Comparative analyses show that the Stacking Ensemble outperforms traditional methods like ANN, RNN, and LSTM, particularly regarding error reduction and computational efficiency. While conventional models often require significant computational resources and are sensitive to noisy data, the Stacking Ensemble demonstrates greater robustness and reliability across various scenarios. This study highlights the ability of the Stacking Ensemble to achieve an optimal balance between accuracy and efficiency, addressing key challenges in modern electricity load forecasting.

2. Modern machine learning methods

2.1. XGBoost

XGBoost (Extreme Gradient Boosting) is a gradient boosting method optimized for high performance and fast speeds. It builds sequential decision trees, each trying to minimize the remaining errors of the previous tree.

Supposedly we have a dataset with n observations and m characteristics. The goal of XGBoost is to minimize the loss function:

$$L(\phi) = \sum_{i=1}^{n} l\,(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{1}$$

In which:

$y_i, \hat{y}_i$ is a loss function measuring the difference between the actual value and the forecast value $y_i, \hat{y}_i$

Ω is the regularization term for the tree model, which was added to control the complexity of the model and prevent overfitting.Ωf_k

K is the number of decision trees in the model.

f_k is the Kth determinant tree in the model.f_k

In power load forecasting, XGBoost utilizes historical data and factors like time, day, and temperature to build an accurate model. It addresses short- and long-term forecasting challenges, managing non-linear relationships, reducing errors, and preventing overfitting. XGBoost improves forecast accuracy and efficiency across various scenarios.

2.2. LightGBM

LightGBM (Light Gradient Boosting Machine) is an advanced variant of the Gradient Boosting algorithm, specifically designed to handle large and high-dimensional datasets with increased speed and performance. Unlike the traditional "Level-wise" approach, LightGBM uses a "Leaf-wise" technique to build decision trees, significantly enhancing the model's performance by focusing on the most significant splits. While LightGBM is rooted in the Gradient Boosting principle applied to decision trees, it incorporates several optimizations that accelerate training and reduce memory usage, making it highly efficient for large-scale data processing.

Like other Gradient Boosting models, LightGBM's general goal function:

$$\mathrm{L}(\phi) = \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{l}\,(\mathrm{y_i}, \hat{\mathrm{y}}_\mathrm{i}) + \sum_{\mathrm{k}=1}^{\mathrm{K}} \Omega(\mathrm{f_k}) \tag{2}$$

In which:

$y_i, \hat{y}_i$ is the loss function between the actual value and the forecast value

(f_k) is the regularization term that controls the complexity of the tree that determines fkf_kfk.

n is the number of observations in the data.

K is the number of decision trees in the model.

2.3. Random Forest

Random Forest is an ensemble algorithm that uses multiple decision trees to improve prediction accuracy. Known for its robustness, it reduces overfitting and enhances performance in tasks like power load forecasting. The model trains each tree on random subsets of the data, and for regression tasks, such as load forecasting, the final prediction is the average of all tree outputs.

$$\hat{\mathrm{y}} = \frac{1}{\mathrm{M}} \sum_{\mathrm{m}=1}^{\mathrm{M}} \hat{\mathrm{y}}_\mathrm{m} \tag{3}$$

In which:

$\hat{y}$ is the aggregate prediction value.

M is the total number of trees in the forest.

$\hat{y}$ is the predicted value from the mth decision tree

Random Forest is a robust algorithm for short- and long-term load forecasting. It combines multiple decision trees to improve prediction accuracy and reduce overfitting. By averaging outputs, it ensures stable and reliable predictions. The algorithm handles large datasets, manages missing data, and processes non-linear relationships effectively. Its high accuracy makes it ideal for practical applications in the energy sector. Random Forest supports grid management, energy supply optimization, and demand forecasting. It provides deeper insights into factors influencing energy demand, such as weather, time, and seasonal patterns. This flexibility makes it highly valuable for improving decision-making. Organizations can rely on Random Forest to achieve accurate forecasts and ensure operational efficiency in energy-related tasks.

2.4. Support Vector Regression

Support Vector Regression (SVR) is a machine learning method adapted from Support Vector Machines (SVM) principles and explicitly designed for regression tasks. Unlike traditional regression methods, SVR identifies a hyperplane in a high-dimensional space that best predicts the output values. SVR aims to minimize prediction errors while ensuring that most data points lie within a predefined tolerance margin, known as the epsilon margin.

SVR achieves this by balancing two objectives: reducing the error between predicted and actual values and maintaining model simplicity to avoid overfitting. SVR can efficiently handle non-linear relationships between features and the target variable by introducing kernel functions (e.g., linear, polynomial, or radial basis function). This versatility makes SVR particularly useful for tasks involving complex patterns and noisy data, such as load forecasting, stock price prediction, and other regression problems.

2.5. Proposed research methodology

2.5.1. Stacking Ensemble Method

Stacking is an ensemble machine-learning technique that combines predictions from multiple models to generate a final, more accurate prediction. The core idea of stacking is to leverage the strengths of diverse machine learning models to improve forecasting accuracy. In a stacking framework, the initial predictive models, known as base models, are trained on the input data to generate predictions. These predictions are then fed into a secondary model, the meta-model, which learns from them to produce the final forecast.

A significant advantage of stacking is its flexibility in combining various models, such as linear regression, Random Forest, XGBoost, LightGBM, SVR, ANN, CNN, RNN, and LSTM, to create a more robust model with higher generalization capabilities and reduced forecast errors. This study selected four independent machine learning models—XGBoost, LightGBM, Random Forest, and SVR—as base models due to their complementary strengths in handling complex data patterns. These base models were trained on the same dataset (X_train, y_train) and, after training, generated predictions for the test dataset (X_test). These predictions were then combined to form a new dataset (X_meta), which served as input for the meta-model. The meta-model, often a simple algorithm such as linear regression (Ridge), was trained on this new dataset (X_meta) using the actual results (y_test) as the target output. Once trained, the meta-model integrated the predictions from the base models to produce the final forecast. This two-level architecture not only enhances accuracy but also improves stability by addressing the weaknesses of individual models.

To evaluate the performance of the Stacking Ensemble model, standard error metrics such as MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), MSE (Mean Squared Error), and RMSE (Root Mean Squared Error) were used. The results of the base models and the Stacking Ensemble were compared with the actual values and visualized to clearly illustrate the differences in accuracy and overall performance.

2.5.2. Algorithm Flowchart

Figure 1 illustrates the load forecasting data preparation process, starting with raw data collection from sources like databases or APIs. The data is preprocessed through cleaning, normalization, and encoding, resulting in a clean dataset. Features (X) and labels (y) are split, then divided into training (X_train, y_train) and testing sets (X_test, y_test). The training set trains the model, while the testing set evaluates its accuracy and reliability for real-world predictions.

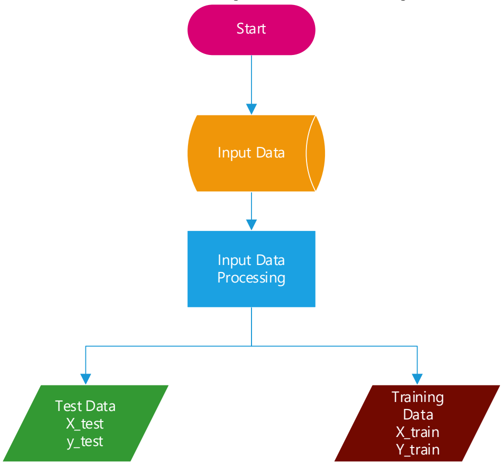

Figure 1. *Input data processing algorithm flowchart*

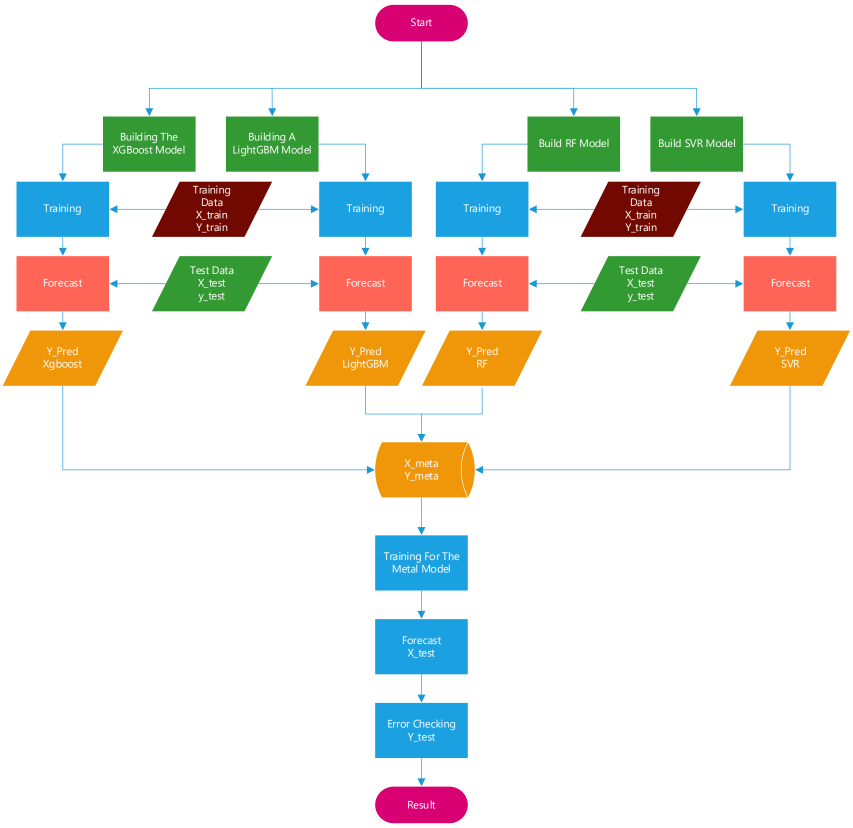

Figure 2. *Algorithmic flowchart of the Stacking model*

Figure 2 illustrates constructing a load forecasting model using multiple machine learning techniques to optimize accuracy. Raw data is collected from sources like databases, files, or APIs and processed into training (X_train, y_train) and testing sets (X_test, y_test). Four models—XGBoost, LightGBM, Random Forest (RF), and Support Vector Regression (SVR)—are trained separately on the training set. Each model generates predictions (Y_pred) on the test set, which are combined into a new dataset (X_meta, y_meta) to train a meta-model. The meta-model acts as an ensemble layer, integrating the base models' outputs for more accurate predictions. After evaluation on a new test set, the meta-model provides the final forecast, leveraging the strengths of each base model to improve accuracy and reliability in load forecasting.

3. Results and Discussion

3.1. Electrical Load Data

The load data utilized in this study comprises the electricity load in Ho Chi Minh City, as presented in Table 1, recorded from 2010 to 2018, with measurements taken at hourly intervals. Table 1 details the key parameters and characteristics of this dataset. This dataset will be divided into two subsets, training and testing data, to facilitate forecasting electricity load in Ho Chi Minh City. The training data will be used to build and fine-tune the forecasting models. In contrast, the testing data will be employed to evaluate the models' accuracy and performance in predicting future electricity demand.

Table 1. *Input Load Data*

Date\Time	0:00 hours	....	23:00 hours
10/11/2010	1117.4 MW	....	1305.5 MW
10/13/2010	1298.3 MW	....	1278.9 MW
....	....	....	....
31/12/2018	1902.7 MW	....	2059.5 MW

3.2. Results

Figures 3, 4, 5, 6, and 7 present graphs of predicted and actual values for the Random Forest, SVR, XGBoost, LightGBM, and Stacking models, respectively. The graphs show that the predicted values closely align with the actual values, demonstrating the effectiveness of the applied models.

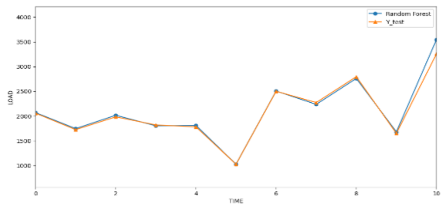

Figure 3. *The graph of predicted values using the Random Forest model*

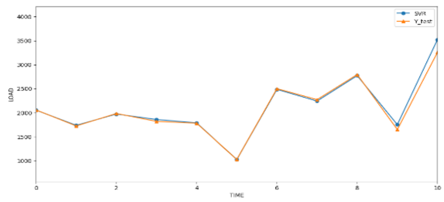

Figure 4. *The graph of predicted values using the SVR model*

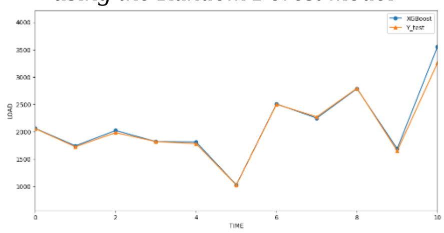

Figure 5. *The graph of predicted values using the XGBoost model*

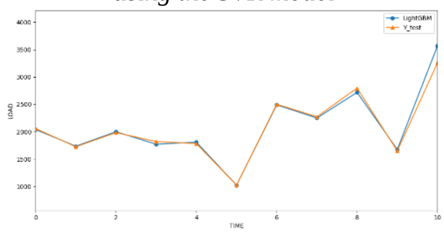

Figure 6. *The graph of predicted values using the LightGBM model*

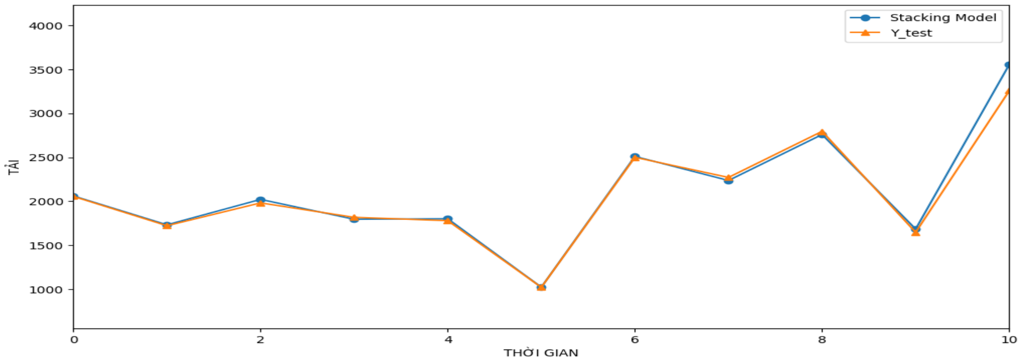

Figure 7. *The graph of predicted values using the Stacking model*

Table 2 summarizes the error metrics (MAE, MAPE, MSE, and RMSE) for the Random Forest, SVR, XGBoost, LightGBM, and Stacking models. The analysis shows that the Stacking model outperforms the others, achieving the lowest error values, including an MAE of 30.18, MAPE of 1.35%, MSE of 1912.63, and RMSE of 43.73. These results indicate that Stacking provides the highest predictive accuracy and the most effective error reduction, making it the optimal model for this task. In contrast, LightGBM recorded the highest error values across all metrics, with an MAE of 37.15, MAPE of 1.68%, MSE of 2617.46, and RMSE of 51.16. This demonstrates that LightGBM is significantly less effective in load forecasting. XGBoost performed better than LightGBM, with an MAE of 31.39, MAPE of 1.41%, MSE of 2142.72, and RMSE of 46.29, indicating a moderate performance. Similarly, Random Forest achieved comparable results with an MAE of 30.43, MAPE of 1.36%, MSE of 2007.99, and RMSE of 44.81, proving it to be a reliable forecasting model. SVR showed average performance, with an MAE of 34.43, MAPE of 1.56%, MSE of 2287.92, and RMSE of 47.83, ranking below Stacking but above LightGBM. Overall, Stacking outperforms the other models by effectively combining multiple base models, delivering higher predictive accuracy and better adaptability to data variability. This makes Stacking the most optimal choice among the applied models.

Table 2. *Error values of the models*

Model	MAE	MAPE	MSE	RMSE
Random Forest	30.42962849	0.013608948	2007.986	44.81056
SVR	34.42934989	0.015570728	2287.916	47.83216
XGBoost	31.38944068	0.014104695	2142.716	46.28948
LightGBM	37.1546672	0.016842936	2617.463	51.16114
Stacking	30.17780634	0.013536993	1912.633	43.73366

4. Conclusion

The Stacking model, combining base models like Random Forest, SVR, XGBoost, and LightGBM, outperforms individual models in short-term load forecasting by achieving lower MAE, MAPE, MSE, and RMSE values. It mitigates base model limitations, such as the higher errors of LightGBM, by integrating outputs to enhance accuracy and adaptability.

Beyond load forecasting, this research aligns with advancements in machine learning across industries, showcasing the adaptability of Stacking beyond energy applications. Expanding to fields like weather forecasting and finance enhances diversity. Weather datasets involve complex factors like temperature and humidity, demanding high-accuracy models. In finance, where data is non-linear and volatile, such as stock price fluctuations, Stacking leverages base model strengths for optimized predictions. Real-world validation in these areas will strengthen its practical value.

Broadening applications test the flexibility of Stacking across datasets, reinforcing its robustness. Future work may refine base model parameters, explore algorithms like CatBoost or Neural Networks, and investigate ensemble methods like Blending or Bagging. Enhancing data processing, feature extraction, and integrating real-time data can improve adaptability. Stacking can also optimize energy distribution in smart grids or EMS, leading to cost savings, efficiency, and environmental benefits. Expanding its role in decision-support systems across sectors will improve resource management and decision-making, reinforcing its real-world impact.

REFERENCES

[1] A. V. Ves, N. Ghitescu, C. Pop, M. Antal, T. Cioara, I. Anghel, and I. Salomie, "A Stacking Multi-Learning Ensemble Model for Predicting Near Real Time Energy Consumption Demand of Residential Buildings," *Proc. - 2019 IEEE 15th Int. Conf. Intell. Comput. Commun. Process,* 2019, pp. 183-189, doi: 10.1109/ICCP48234.2019.8959572.

[2] H. M. Al-Hamadi, "Long-term electric power load forecasting using fuzzy linear regression technique," *PEAM 2011 - Proc. 2011 IEEE Power Eng. Autom. Conf.*, vol. 3, pp. 96-99, 2011, doi: 10.1109/PEAM.2011.6135023.

[3] S. Chen, R. Lin, and W. Zeng, "Short-Term Load Forecasting Method Based on ARIMA and LSTM," *Int. Conf. Commun. Technol. Proceedings, ICCT*, 2022, pp. 1913-1917, doi: 10.1109/ICCT56141.2022.10073051.

[4] M. M. Eljazzar and E. E. Hemayed, "Enhancing electric load forecasting of ARIMA and ANN using adaptive Fourier series," *the 7th of 2017. Comput. Commune. Work. Conf. CCVC 2017*, January 2017, doi: 10.1109/CCWC.2017.7868457.

[5] H. Yiling and H. Shaofeng, "A Short-Term Load Forecasting Model Based on Improved Random Forest Algorithm," *Proc. - 2020 7th Int. Forum Electr. Eng. Autom. IFEEA*, 2020, no. 1, pp. 928-931, doi: 10.1109/IFEEA51475.2020.00195.

[6] C. Wang, J. Zhang, L. Tian, L. Xue, Y. Zheng, and L. Liu, "Short-term Load Forecasting Based on Kprototypes Clustering and Random Forest," *5th IEEE Conf. Energy Internet Energy Syst. Integr. Energy Internet Carbon Neutrality,* 2021, pp. 1226-1230, doi: 10.1109/EI252483.2021.9712977.

[7] P. F. Pai and W. C. Hong, "Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms," *Electr. Power Syst. Res.*, vol. 74, no. 3, pp. 417-425, 2005, doi: 10.1016/j.epsr.2005.01.006.

[8] L. A. Abad, S. M. Sarabia, J. M. Yuzon, and M. C. Pacis, "A Short-Term Load Forecasting Algorithm Using Support Vector Regression & Artificial Neural Network Method (SVR-ANN)," *2020 11th IEEE Control Syst. Grad. Res. Colloquium*, 2020, pp. 138-143, doi: 10.1109/ICSGRC49013.2020.9232630.

[9] H. Ma, P. Yang, F. Wang, X. Wang, D. Yang, and B. Feng, "Short-Term Heavy Overload Forecasting of Public Transformers Based on Combined LSTM-XGBoost Model," *Energies*, vol. 16, no. 3, 2023, doi:10.3390/EN16031507.

[10] M. Mohamed, F. E. Mahmood, M. A. Abd, M. Rezkallah, and A. Hamadi, "Load Demand Forecasting Using eXtreme Gradient Boosting (XGBoost)," *2023 Ei Ind. APPL. Soc. Then. Meet.*, 2023, pp. 1-7, doi: 10.1109/IAS54024.2023.10406613.

[11] H. Auto-tuning, S. Lei, and F. Wang, "A Short-term Net Load Forecasting Method Based on Two-stage Feature Selection and LightGBM with," *2023 IEEE/IAS 59th Ind. Commer. Power Syst. Tech. Conf.*, 2023, pp. 1-6, doi: 10.1109/ICPS57144.2023.10142095.

[12] Z. Fang, J. Zhan, J. Cao, L. Gan, and H. Wang, "Research on Short-Term and Medium-Term Power Load Forecasting Based on STL-LightGBM," *2nd Int. Conf. Electr. Eng. Control Sci.*, 2022, pp. 1047-1051, doi: 10.1109/IC2ECS57645.2022.10088145.

[13] U. Samal and A. Kumar, "Enhancing Software Reliability Forecasting Through a Hybrid ARIMA-ANN Model," *Arab. J. Sci. Eng.*, vol. 48, 2023, doi: 10.1007/s13369-023-08486-1.

[14] B. Satish, K. S. Swarup, S. Srinivas, and A. H. Rao, "Effect of temperature on short-term load forecasting using an integrated ANN," *Electr. Power Syst. Res.*, vol. 72, no. 1, pp. 95-101, 2004, doi: 10.1016/j.epsr.2004.03.006.

[15] M. Alhussein, K. Aurangzeb, and S. I. Haider, "Hybrid CNN-LSTM model for short-term individual household load forecasting," *IEEE Access*, vol. 8, pp. 180544-180557, 2020.

[16] W. Xiong, "Bus load forecasting based on maximum information coefficient and CNN-LSTM model,"

IEEE Int. Conf. Image Process. Comput. Appl., 2023, pp. 659-663, doi: 10.1109/ICIPCA59209.2023.10257944.

[17] A. Ajitha, M. Goel, M. Assudani, S. Radhika, and S. Goel, "Design and development of Residential Sector Load Prediction model during COVID-19 Pandemic using LSTM based RNN," *Electr. Power Syst. Res.*, vol. 212, 2022, Art. no. 108635, doi: 10.1016/j.epsr.2022.108635.

[18] W. Xiangxue, X. Lunhui, and C. Kaixun, "Data-Driven Short-Term Forecasting for Urban Road Network Traffic Based on Data Processing and LSTM-RNN," *Arab. J. Sci. Eng.*, vol. 44, pp. 3043-3060, 2018, doi: 10.1007/s13369-018-3390-0.

[19] X. Chen, M. Yang, Y. Zhang, J. Liu, and S. Yin, "Load Prediction Model of Integrated Energy System Based on CNN-LSTM," *3rd Int. Conf. Energy Eng. Power Syst.*, 2023, pp. 248-251, doi: 10.1109/EEPS58791.2023.10257124.

[20] X. Guo, Q. Zhao, D. Zheng, Y. Ning, and Y. Gao, "A short-term load forecasting model of multi-scale CNN-LSTM hybrid neural network considering the real-time electricity price," *Energy Reports*, vol. 6, pp. 1046-1053, 2020, doi: 10.1016/j.egyr.2020.11.078.

[21] Z. He, S. Chen, N. Pan, T. Ba, C. Lin, X. Yang, and G. Li, "Short-Term Power Load Forecasting of Multi-Charging Piles Based on Improved Gate Recurrent Unit," *IEEE Access*, vol. 10, pp. 2490-2499, 2024.

[22] C. Time-series, I. E. Livieris, E. Pintelas, S. Stavroyiannis, and P. Pintelas, "Ensemble Deep Learning Models for Forecasting," *Algorithms*, vol. 13, pp. 1-21, 2020, doi:10.3390/a13050121.

[23] S. Singh and M. M. Tripathi, "A Comparative Analysis of Extreme Gradient Boosting Technique with Long Short-Term Memory and Layered Recurrent Neural Network for Electricity Demand Forecast," *6th Int. Conf. Recent Trends Electron. Information, Commun. Technol.*, 2021, pp. 297-302, doi: 10.1109/RTEICT52294.2021.9573988.

[24] M. Goyal and M. Pandey, "A Systematic Analysis for Energy Performance Predictions in Residential Buildings Using Ensemble Learning," *Arab. J. Sci. Eng.*, vol. 46, 2020, doi: 10.1007/s13369-020-05069-2.

[25] L. Zhang and D. Jánošík, "Enhanced short-term load forecasting with hybrid machine learning models: CatBoost and XGBoost approaches," *Expert Syst. Appl.*, vol. 241, 2024, doi: 10.1016/j.Eshwa.2023.122686.

[26] M. Ulagammai, "Short Term Load Forecasting Using ANN and WNN," *Proc. Int. Conf. Intell. Innov. Technol. Comput. Electr. Electron. ICIITCEE*, 2023, pp. 612-616, doi: 10.1109/IITCEE57236. 2023.10091081.

[27] M. Abumohsen, A. Y. Owda, and M. Owda, "Electrical Load Forecasting Using LSTM, GRU, and RNN Algorithms," *Energies*, vol. 16, no. 5, pp. 1-31, 2023, doi: 10.3390/en16052283.

[28] J. Zhang, Z. Zhu, and Y. Yang, "Electricity Load Forecasting Based on CNN-LSTM," *IEEE Int. Conf. Electr. Autom. Comput. Eng.*, 2023, pp. 1385-1390, doi: 10.1109/ICEACE60673. 2023.10442217.

[29] M. Dostmohammadi, M. Z. Pedram, S. Hoseinzadeh, and D. A. Garcia, "A GA-stacking ensemble approach for forecasting energy consumption in a smart household: A comparative study of ensemble methods," *J. Environ. Manage.*, vol. 364, 2024, Art. no. 121264, doi: 10.1016/j.genvman.2024.121264.

[30] A. Ghasemieh, A. Lloyd, P. Bahrami, P. Vajar, and R. Kashef, "A novel machine learning model with Stacking Ensemble Learner for predicting emergency readmission of heart-disease patients," *Decis. Anal. J.*, vol. 7, 2023, Art. no. 100242, doi: 10.1016/j.dajour.2023.100242.