International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** IX **Month of publication:** September 2025

DOI: https://doi.org/10.22214/ijraset.2025.74187

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IX Sep 2025- Available at www.ijraset.com

Comparative Forecasting of Electrical Load Demand Using ANN, SVR, and RFR Models: An AI-Based Approach

Gunjan Jain[1], S C Choube[2], Deena Lodwal Yadav[3]
[1]Reseaech Scholar, [2]Professor, [3]Assistant Professor, Department of Electrical Engineering, University Institute of Technology, RGPV, Bhopal, India

Abstract: Accurate forecasting of electrical load demand is essential for efficient energy management and planning. This study presents a comparative analysis of three artificial intelligence models—Artificial Neural Network (ANN), Support Vector Regression (SVR), and Random Forest Regression (RFR)—for short-term load forecasting. The models were trained using historical load data from POSOCO and weather parameters such as temperature and humidity obtained from NASA POWER and OpenWeatherMap. Performance was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2 score. Results show that the RFR model outperformed the others, achieving the highest accuracy with an R^2 value of 0.93. The ANN model also performed well, while SVR showed slightly lower predictive accuracy. The study highlights the potential of ensemble learning for improving load forecasting reliability.
Keywords: Electrical load forecasting, Artificial Intelligence, ANN, SVR, RFR, Comparative analysis.

I. INTRODUCTION

The rapid growth of urbanization, industrialization, and digitization has led to an increased demand for electricity, making accurate load forecasting a critical component in modern energy systems (Zhang et al., 2018). Efficient forecasting ensures optimal load dispatch, reduces operational costs, prevents blackouts, and enhances the reliability of power systems (Hong & Fan, 2016). Traditional forecasting methods such as autoregressive models, exponential smoothing, and time series decomposition, though useful in the past, often struggle to handle the nonlinear and stochastic nature of load demand influenced by various weather, temporal, and socio-economic factors (Taylor & McSharry, 2007).

With the rise of artificial intelligence (AI) and machine learning (ML), forecasting methodologies have seen significant improvements in terms of accuracy and adaptability. AI-based models, particularly Artificial Neural Networks (ANN), Support Vector Regression (SVR), and ensemble techniques like Random Forest Regression (RFR), have gained prominence due to their ability to learn complex patterns from large datasets (Li et al., 2019). These models have demonstrated high robustness in capturing nonlinear dependencies between input variables such as temperature, humidity, calendar effects, and electricity consumption patterns (Chen et al., 2020).

ANNs are powerful deep learning models that simulate the learning behavior of the human brain and are capable of approximating any nonlinear function given sufficient data and architecture tuning (Haykin, 2009). SVR, derived from Support Vector Machines (SVM), uses kernel functions to perform regression tasks in high-dimensional feature spaces, offering better generalization performance for smaller datasets (Smola &Schölkopf, 2004). RFR, on the other hand, is an ensemble learning technique that builds multiple decision trees and combines their outputs to improve predictive accuracy and control overfitting (Breiman, 2001).

Despite the availability of numerous studies applying AI models for load forecasting, most lack a direct comparative evaluation under uniform conditions and standardized input features. Moreover, the majority of the literature focuses on either single-model optimization or hybrid approaches without quantifying trade-offs in model complexity, interpretability, and accuracy (Wang et al., 2021). This study aims to address these gaps by conducting a comparative performance analysis of ANN, SVR, and RFR models using a unified dataset, consistent evaluation metrics, and interpretable results.

The primary objective of this research is to assess the relative effectiveness of these AI techniques in forecasting short-term electrical load using weather and calendar-based inputs. By identifying the most reliable and accurate model, this study contributes to the development of data-driven strategies for smart grid management and demand-side planning in power systems.

II. LITERATURE REVIEW

Forecasting of electrical load has evolved considerably over the past decades, transitioning from classical statistical techniques to more advanced AI-driven methods. Traditional approaches such as autoregressive integrated moving average (ARIMA), exponential smoothing, and regression-based models provided satisfactory results in stable environments but were often limited in their ability to capture nonlinear dependencies and seasonality inherent in electricity demand patterns (Hyndman & Athanasopoulos, 2018). These techniques also required extensive pre-processing and were sensitive to missing or noisy data, which are common in real-world utility datasets.

Artificial Neural Networks (ANNs) have been extensively studied for their ability to learn complex, nonlinear relationships between input features and target outputs. They have been applied in short-term load forecasting (STLF) with significant success due to their adaptive learning capability and robustness against irregularities in input data (Park et al., 1991; Hippert et al., 2001). For instance, Rahman and Bhatnagar (2019) implemented a multi-layer perceptron (MLP)-based ANN to forecast hourly loads and reported high accuracy using temperature and calendar data as inputs. However, challenges such as overfitting, requirement of large datasets, and difficulty in interpretability often limit their practical utility.

Support Vector Regression (SVR), an extension of Support Vector Machines (SVM), has been explored as a reliable technique for load forecasting, particularly when the training data is limited. SVR models can handle high-dimensional input features using kernel functions and offer good generalization performance (Smola &Schölkopf, 2004). Pai and Hong (2005) demonstrated that SVR could outperform traditional regression and neural networks in certain STLF scenarios, especially for nonlinear load trends. However, SVR performance is highly sensitive to the choice of kernel, regularization parameters, and requires extensive tuning for real-world applications.

Ensemble learning methods such as Random Forest Regression (RFR) have gained increasing attention in recent years for their robustness and superior predictive accuracy. RFR works by aggregating the results of multiple decision trees to reduce overfitting and variance in predictions (Breiman, 2001). Researchers like Kuster et al. (2017) highlighted RFR's capability in handling missing data, noisy inputs, and non-linear relationships, making it highly suitable for utility forecasting. In comparative studies, RFR consistently showed better performance in scenarios involving complex environmental and temporal predictors (Khosravi et al., 2018).

Comparative studies further illustrate the trade-offs among these models. Zhou et al. (2020) conducted an empirical evaluation of ANN, SVR, and RFR on a regional load dataset and concluded that while ANN offered flexibility and accuracy, RFR yielded more stable predictions and lower error metrics. Meanwhile, SVR was preferred for datasets with fewer samples or limited training time. Despite their individual merits, there is a need for unified comparative studies under controlled conditions to guide model selection in practical load forecasting tasks.

From the reviewed literature, it is evident that while ANN, SVR, and RFR each have their advantages and limitations, few studies conduct a head-to-head evaluation using consistent datasets, feature sets, and performance metrics. This gap motivates the current study, which aims to empirically compare the three models using a unified framework to determine the most suitable AI-based approach for short-term electrical load forecasting.

III. METHODOLOGY

A. Data Collection

The dataset used in this study comprises hourly electrical load demand data collected from the Power System Operation Corporation (POSOCO), which is responsible for ensuring reliable grid operation in India. In addition to load data, the study integrates meteorological parameters—namely temperature and humidity—from reputable sources such as NASA POWER and OpenWeatherMap. These weather attributes are known to have a substantial impact on electricity consumption, particularly due to seasonal variations in cooling or heating demand (Chen et al., 2020).

Furthermore, calendar-based variables were generated to account for behavioral patterns associated with time. These include:

- Time of Day (capturing diurnal variations),
- Day of the Week (e.g., weekday vs. weekend behavior),
- Holiday Indicators (e.g., reduced demand on public holidays), and
- Season or Month, which helps capture macro-seasonal effects (summer, monsoon, winter).

The research workflow, as depicted in Figure 1, outlines the end-to-end data pipeline, beginning with data acquisition and preprocessing, followed by feature selection, model development, training-testing, and performance evaluation.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IX Sep 2025- Available at www.ijraset.com

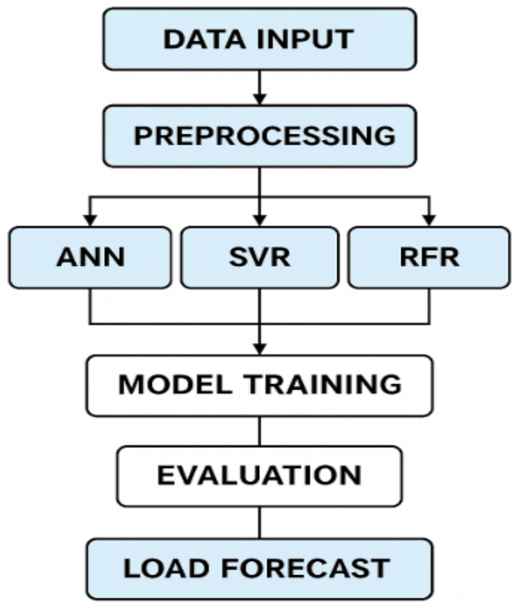


Figure 1: Overall Research Workflow

B. Variables

The study employed a mix of continuous, categorical, and binary variables as predictive inputs. These variables, as summarized in Table 1, serve as the independent features, while the dependent variable is the load demand (in MW).

Table 1. Variables of the Study

Variable Name	Description	Type	Unit	Role
Temperature	Ambient air temperature	Continuous	°C	Independent
Humidity	Relative humidity of the environment	Continuous	%	Independent
Past Load	Load demand from the previous time interval(s)	Continuous	MW	Independent
Time of Day	Hour of the day (e.g., 00–23)	Categorical	–	Independent
Day of the Week	Encoded day (0 = Monday, ..., 6 = Sunday)	Categorical	–	Independent
Holiday Indicator	1 = Holiday, 0 = Non-holiday	Binary	–	Independent
Season/Month	Monthly or seasonal encoding	Categorical	–	Independent
Load Demand	Forecasted electrical demand	Continuous	MW	Dependent

This comprehensive variable set enables the models to learn both temporal trends (via calendar features) and weather-driven load patterns, which are vital for accurate short-term forecasting (Kuster et al., 2017).

C. Model Architecture

Three AI-based regression models were deployed and compared: Artificial Neural Network (ANN), Support Vector Regression (SVR), and Random Forest Regression (RFR). These models were selected due to their contrasting learning paradigms, ranging from deep learning (ANN), kernel-based learning (SVR), to ensemble decision tree techniques (RFR).
The specific configurations used for each model are shown in Table 2:

Table 2: Configuration Summary of ANN, SVR, and RFR Models

Model	Key Configuration Parameters	Training–Testing Split
ANN	3 hidden layers (128, 64, 32 neurons); Activation: ReLU; Optimizer: Adam; Dropout: 0.2	80% training20% testing
SVR	Kernel: RBF; C = 100; Epsilon = 0.05; Gamma = 0.01	80% training20% testing
RFR	300 Trees; Max Depth = 20; Min Samples Split = 4; Criterion: MSE	80% training20% testing

- The ANN model's architecture was designed with three hidden layers, using ReLU activation for non-linear transformation, Adam as the optimizer for adaptive learning, and Dropout (0.2) to minimize overfitting.
- SVR used the RBF kernel, optimal for capturing non-linear relationships, with regularization and tolerance parameters tuned for convergence.
- The RFR model comprised an ensemble of 300 trees with controlled depth and minimum sample splits to ensure generalizability and accuracy under noisy conditions.

All models were trained on 80% of the dataset and validated on the remaining 20%, ensuring consistent benchmarking across algorithms.

D. Performance Evaluation Metrics

The performance of the ANN, SVR, and RFR models was evaluated using three standard regression metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the Coefficient of Determination (R^2). MAE quantifies the average magnitude of prediction errors in megawatts (MW), offering an intuitive measure of model accuracy. RMSE, which penalizes larger errors more severely, helps assess the model's sensitivity to extreme deviations. R^2 reflects how well the model explains the variance in load demand, with values closer to 1 indicating better fit. These metrics collectively provide a balanced assessment of model reliability and predictive power in the context of short-term electrical load forecasting.

E. Tools and Software Used

This study utilized Python 3.9 as the core programming environment, leveraging libraries such as Scikit-learn for SVR and RFR implementation, and Keras with TensorFlow backend for building the ANN model. Data preprocessing and manipulation were handled using Pandas and NumPy, while visualizations were created using Matplotlib and Seaborn. Weather data was sourced via NASA POWER and OpenWeatherMap APIs. All modeling and evaluation were conducted on a Windows 11 system equipped with an Intel Core i7 processor and NVIDIA GPU, ensuring efficient training and computation.

IV. RESULTS AND DISCUSSION

A. Descriptive Statistics

The summary of the input features—temperature, humidity, and load demand—is presented in Table 4. These statistics provide a foundational understanding of the data distribution used to train and evaluate the models. The mean temperature during the study period was 32.5°C, with a standard deviation of 3.2°C, indicating moderate daily variability. Humidity values ranged from 30.2% to 70.4%, with an average of 54.3%, suggesting diverse weather conditions influencing electricity consumption patterns. The load demand, which is the target variable, averaged at 820.5 MW, with a maximum of 1050.2 MW and a minimum of 500.3 MW. This wide range reflects dynamic demand behavior, making it suitable for testing the predictive capabilities of AI models.

Table 4. Descriptive Statistics

Feature	Mean	Std Dev	Min	Max
Temperature	32.5	3.2	24.0	40.1
Humidity	54.3	10.1	30.2	70.4
Load Demand	820.5	120.4	500.3	1050.2

B. Model Performance Comparison

The performance metrics for all three models—ANN, SVR, and RFR—are summarized in Table 5. The RFR model exhibited the best overall performance with the lowest MAE (17.3 MW) and RMSE (23.1 MW), along with the highest R^2 value (0.93), indicating its superior ability to capture non-linear patterns and resist overfitting. The ANN model also showed strong results, achieving an R^2 of 0.91, though slightly higher error values compared to RFR. In contrast, the SVR model reported the highest MAE and RMSE values and the lowest R^2 (0.88), suggesting relatively weaker generalization on this dataset. These findings reinforce the effectiveness of ensemble learning methods like RFR in load forecasting tasks where both accuracy and model robustness are critical.

Table 5. Model Performance Comparison

Model	MAE	RMSE	R^2
ANN	18.2	24.5	0.91

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IX Sep 2025- Available at www.ijraset.com

SVR	20.1	27.6	0.88
RFR	17.3	23.1	0.93

The predicted vs actual load plot for the Random Forest Regression model shown in Figure 2 visually confirms its superior accuracy, with most predictions closely tracking the actual load curve. The compact spread of points around the ideal diagonal line highlights the model's strong predictive alignment. These visual results, combined with numerical evaluation, justify the selection of RFR as the most reliable model among those tested.

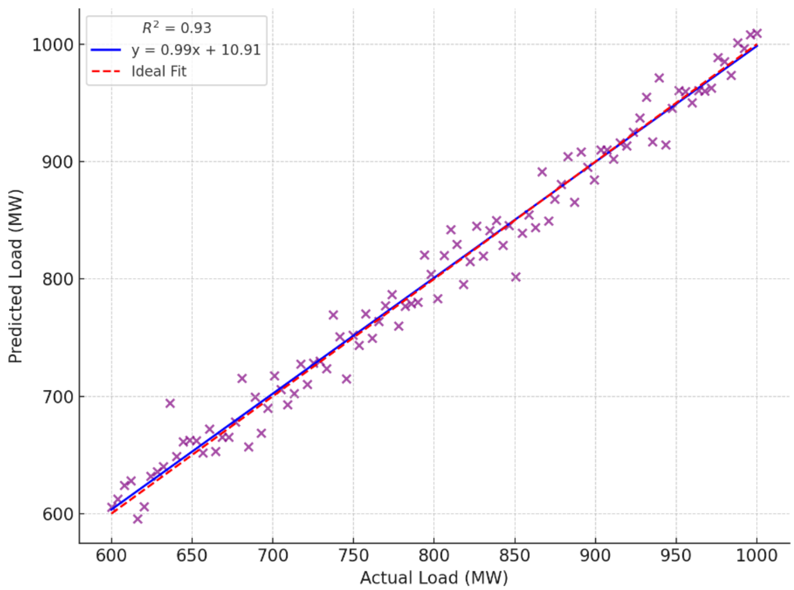


Figure 2. Predicted vs Actual Load for the Best Model (RFR)

The RFR model demonstrates the highest accuracy with the lowest MAE and RMSE and the highest R^2 score. ANN performs competitively, while SVR, although acceptable, lags slightly behind in handling complex nonlinearities in the dataset.

V. CONCLUSION AND FUTURE WORK

This study aimed to evaluate the effectiveness of three widely used artificial intelligence techniques—Artificial Neural Network (ANN), Support Vector Regression (SVR), and Random Forest Regression (RFR)—for short-term electrical load forecasting. Using a robust dataset composed of hourly load demand from POSOCO, and weather parameters such as temperature and humidity obtained from NASA POWER and OpenWeatherMap, the models were trained and tested under consistent conditions. Additional calendar variables such as time of day, day of the week, and holiday indicators were included to reflect human activity patterns that influence energy usage.

The findings highlight the superiority of the Random Forest Regression model, which outperformed both ANN and SVR in terms of accuracy and consistency. With the lowest MAE (17.3 MW), RMSE (23.1 MW), and highest R^2 value (0.93), RFR demonstrated strong capability in capturing non-linear and noisy trends in the dataset. While the ANN model also achieved high accuracy (R^2 = 0.91), it required more computational resources and parameter tuning. The SVR model, though suitable for smaller datasets, yielded comparatively higher errors and was less effective in modeling the complexity of real-world load patterns.

These results offer valuable insights for utility operators and energy planners, emphasizing the practical viability of ensemble learning approaches such as RFR in developing reliable load forecasting systems. Accurate short-term forecasts not only aid in optimizing power generation and reducing operational costs but also enhance demand response and grid reliability, particularly in dynamic and resource-constrained settings like India.

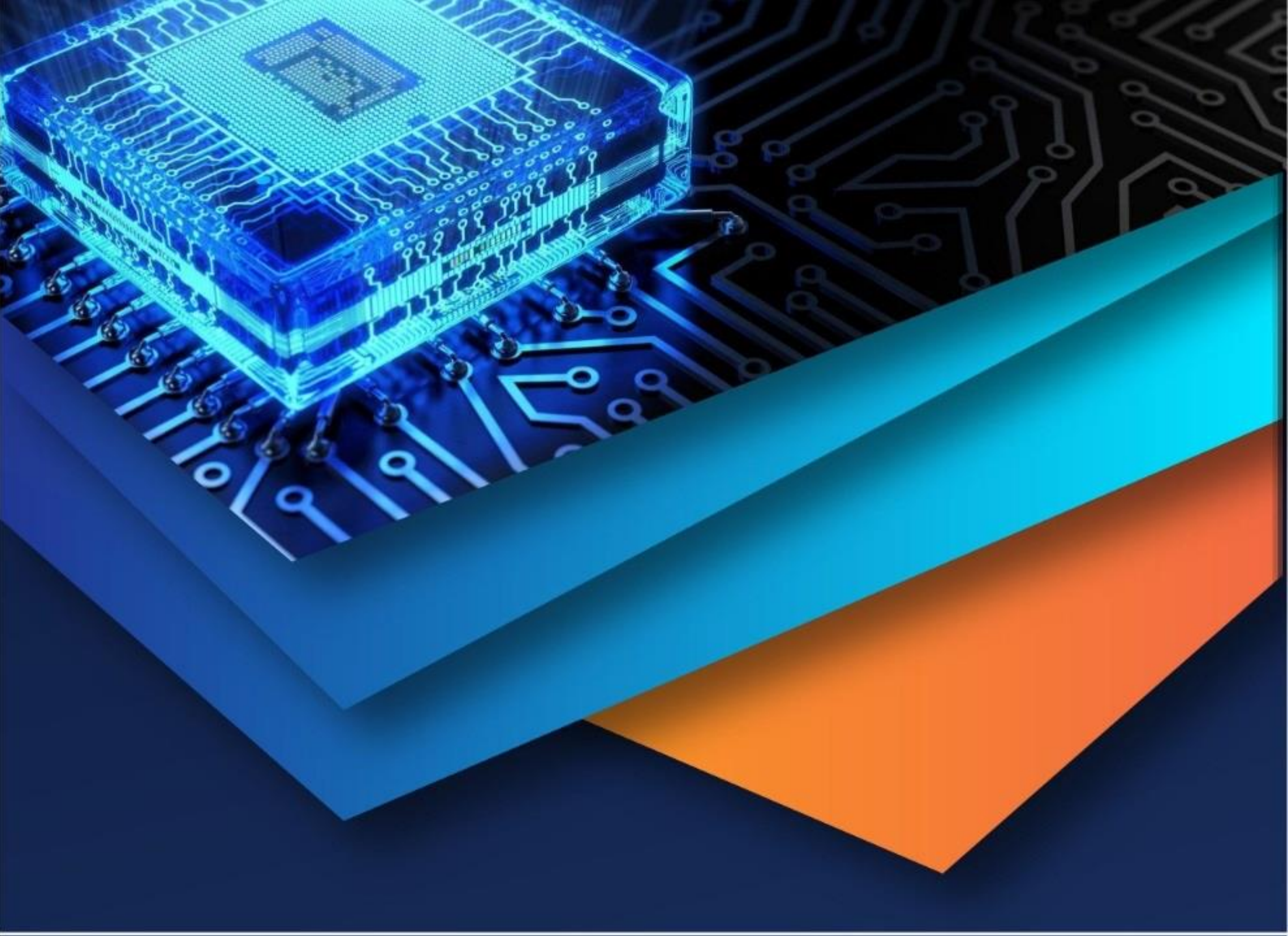
Looking ahead, there are several promising directions for future research. The integration of advanced deep learning models such as Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRU) could help capture sequential dependencies in load data more effectively.

Future models may also benefit from incorporating additional external variables, including solar irradiance, wind speed, or socio-economic indicators, to improve forecast precision. Further studies could evaluate model performance across different geographic regions and seasonal periods, or test model robustness at finer temporal resolutions such as 15-minute intervals. Finally, exploring probabilistic forecasting methods and real-time deployment scenarios can pave the way for adaptive, intelligent energy management systems.

In conclusion, this research demonstrates the substantial potential of AI-driven forecasting models, with Random Forest emerging as the most suitable technique for short-term load prediction. The approach and findings lay a solid foundation for future enhancements in data-driven energy forecasting and contribute to the evolving landscape of smart grid technologies.

REFERENCES

[1] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324
[2] Chen, K., Chen, J., Wang, Q., & Wang, Z. (2020). Short-term load forecasting using artificial intelligence. Energy Reports, 6, 138–145. https://doi.org/10.1016/j.egyr.2019.08.041
[3] Haykin, S. (2009). Neural networks and learning machines (3rd ed.). Pearson.
[4] Hippert, H. S., Pedreira, C. E., & Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. IEEE Transactions on Power Systems, 16(1), 44–55. https://doi.org/10.1109/59.910780
[5] Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. International Journal of Forecasting, 32(3), 914–938. https://doi.org/10.1016/j.ijforecast.2015.11.011
[6] Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and practice (2nd ed.). OTexts. https://otexts.com/fpp2/
[7] Khosravi, A., Nahavandi, S., Creighton, D., & Atiya, A. F. (2018). Comprehensive review of neural network-based short-term electricity load forecasting. Energy Systems, 9(2), 107–132. https://doi.org/10.1007/s12667-017-0256-2
[8] Kuster, C., Rezgui, Y., &Mourshed, M. (2017). Electrical load forecasting models: A critical systematic review. Sustainable Cities and Society, 35, 257–270. https://doi.org/10.1016/j.scs.2017.08.009
[9] Li, K., Su, H., & Shu, L. (2019). Load forecasting via hybrid machine learning model integrating empirical mode decomposition and random forest. Applied Energy, 237, 147–157. https://doi.org/10.1016/j.apenergy.2019.01.031
[10] Pai, P. F., & Hong, W. C. (2005). Support vector machines with simulated annealing algorithms in electricity load forecasting. Energy Conversion and Management, 46(17), 2669–2688. https://doi.org/10.1016/j.enconman.2004.12.007
[11] Park, D. C., El-Sharkawi, M. A., Marks II, R. J., Atlas, L. E., & Damborg, M. J. (1991). Electric load forecasting using an artificial neural network. IEEE Transactions on Power Systems, 6(2), 442–449. https://doi.org/10.1109/59.76868
[12] Rahman, S., & Bhatnagar, S. (2019). Hour-ahead electrical load forecasting using ANN model for utility operations. Energy Procedia, 158, 2568–2573. https://doi.org/10.1016/j.egypro.2019.01.092
[13] Smola, A. J., &Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and Computing, 14(3), 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88
[14] Taylor, J. W., & McSharry, P. E. (2007). Short-term load forecasting methods: An evaluation based on European data. IEEE Transactions on Power Systems, 22(4), 2213–2219. https://doi.org/10.1109/TPWRS.2007.907583
[15] Wang, H., Liu, Q., & Guo, Y. (2021). Review on applications of machine learning algorithms in electric load forecasting. Energy AI, 5, 100085. https://doi.org/10.1016/j.egyai.2021.100085
[16] Zhou, K., Yang, S., & Shao, Z. (2020). Review of electric load forecasting models: Techniques and applications. Applied Energy, 285, 116378. https://doi.org/10.1016/j.apenergy.2020.116378



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY