iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78286

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Electricity Demand Forecasting & Grid Planning for India

D. Yasovardhan Reddy[1], M. Keerthi[2], M. Suchitra[3], M. Reddy Sasidhar[4], B. Likitha[5], M. C. Leelu Mamakarni[6]
[1, 2, 3, 4, 5]*Department of Computer Science and Engineering, (Data Science), Sri Venkateswara College of Engineering and Technology, RVS Nagar, Chittoor, India 517127*
[6]*Assistant Professor, Sri Venkateswara College of Engineering And Technology, (Data Science), RVS Nagar, Chittoor 517127*

Abstract: ***Accurate electricity demand forecasting plays a crucial role in efficient power system operation, grid stability, and long-term infrastructure planning. This study proposes a data-driven approach for electricity demand forecasting and grid planning for India using statistical and machine learning techniques. Historical electricity consumption data along with temporal features such as seasonal patterns and time-based indicators are analyzed to identify demand trends.***

The proposed framework integrates multiple forecasting models, including SARIMA, Prophet, and Random Forest Regressor, to capture both linear seasonal patterns and complex non-linear relationships in electricity demand. Feature engineering techniques such as lag features and time-based variables are employed to improve prediction performance. The models are evaluated using standard error metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE).

Experimental results demonstrate that machine learning models, particularly Random Forest, provide improved forecasting accuracy compared to traditional statistical methods. The forecasted demand is further analyzed to support grid planning decisions, including peak load identification and infrastructure expansion strategies. The proposed system provides a scalable solution that can assist energy planners and policymakers in improving electricity distribution efficiency and supporting future smart grid development.

Keywords: Electricity Demand Forecasting, Machine Learning, Random Forest, SARIMA, Prophet Model, Time Series Analysis, Grid Planning, Energy Consumption Prediction.

I. INTRODUCTION

Electricity is one of the most essential resources for economic development and industrial growth. The increasing demand for electricity in India requires efficient management of power generation and distribution systems. Electricity demand forecasting helps power utilities predict future consumption patterns and maintain grid stability.

Accurate demand forecasting enables better planning of electricity generation, load balancing, and infrastructure development. However, electricity demand is influenced by several factors such as seasonal variations, weather conditions, economic activities, and population growth. Traditional statistical models often fail to capture complex nonlinear relationships in electricity demand data.

Recent advances in machine learning and data analytics provide new opportunities to improve electricity demand forecasting. Machine learning models can analyze large volumes of historical data and identify hidden patterns that help improve prediction accuracy.

This paper proposes a forecasting framework using SARIMA, Prophet, and Random Forest algorithms to predict electricity demand and support grid planning decisions.

II. LITERATURE REVIEW

Electricity demand forecasting has been widely studied in power system research. Traditional forecasting models such as ARIMA and SARIMA have been used to capture seasonal patterns in electricity consumption data. These models perform well for linear time-series data but have limitations when dealing with nonlinear relationships.

Recent research has explored machine learning techniques such as Random Forest, Support Vector Machines, and Neural Networks for electricity demand forecasting. These models can capture complex patterns and improve prediction accuracy.

The Prophet forecasting model introduced by Facebook has also gained popularity due to its ability to model trend and seasonality effectively.

However, there is still a need for integrated forecasting frameworks that combine statistical models with machine learning techniques to improve accuracy and support power grid planning.

III. PROPOSED METHODOLOGY

The proposed methodology aims to develop an accurate electricity demand forecasting system using statistical and machine learning techniques. The system analyzes historical electricity consumption data and extracts important time-based features to improve prediction performance.

Initially, the collected electricity demand dataset undergoes data preprocessing, which includes handling missing values, formatting time indices, and cleaning inconsistent records. After preprocessing, feature engineering techniques are applied to generate useful variables such as time-based features (year, month, day, hour) and lag features that capture previous demand patterns.

Next, multiple forecasting models including SARIMA, Prophet, and Random Forest Regressor are implemented to analyze electricity demand patterns. These models are trained using historical data and tested to evaluate their prediction performance.

Finally, the models are evaluated using performance metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error
(MAPE). The best-performing model is selected to generate electricity demand forecasts, which can support grid planning and energy management decisions.

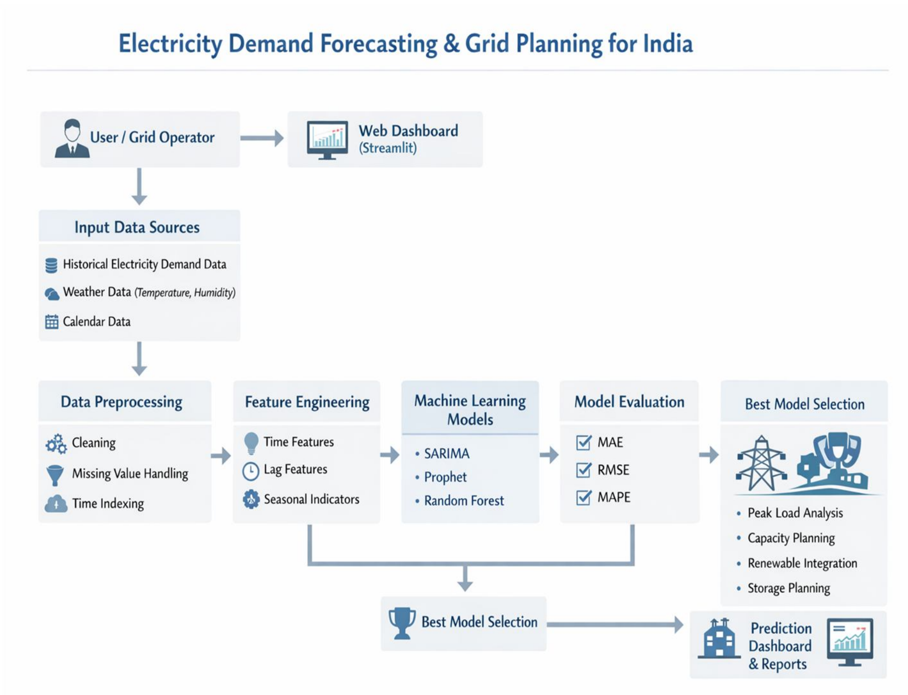


Figure 1: Electricity Demand Forecasting Workflow

IV. SYSTEM ARCHITECTURE

The proposed electricity demand forecasting system follows a structured architecture that processes historical electricity consumption data and generates accurate demand predictions for grid planning.

The architecture begins with the data collection layer, where historical electricity demand datasets and related temporal information are gathered. The collected data is then passed to the data preprocessing module, which handles missing values, converts datetime formats, and prepares the dataset for analysis.

Next, the system performs feature engineering, where useful features such as year, month, day, hour, and lag variables are generated to capture time-based patterns in electricity consumption.

The processed dataset is then used in the model training layer, where forecasting algorithms including SARIMA, Prophet, and Random Forest Regressor are implemented to learn patterns in electricity demand.

After training, the models are evaluated using performance metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) to determine the most accurate model.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

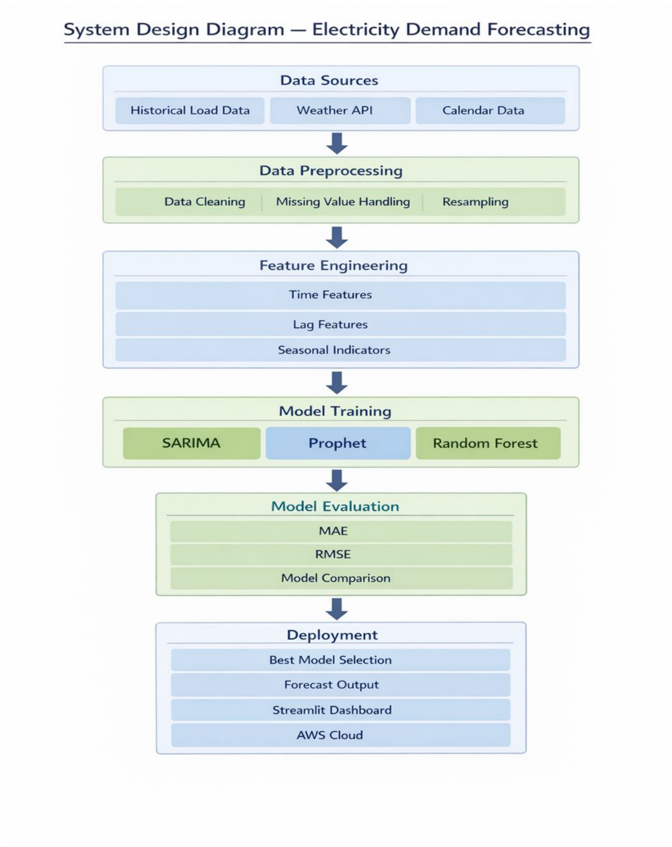


Figure 2: System Block Diagram

V. MACHINE LEARNING MODELS

In this project, multiple statistical and machine learning models are used to forecast electricity demand. These models analyze histor ical electricity consumption data and identify patterns to generate accurate predictions.

A. SARIMA (Seasonal AutoRegressive Integrated Moving Average)

SARIMA is a statistical time-series forecasting model that captures both trend and seasonal patterns in electricity demand data. It is effective for modeling seasonal electricity consumption patterns.

B. Prophet Model

Prophet is an additive time-series forecasting model developed for handling trend, seasonality, and holiday effects in time-series dat a. It provides reliable predictions even when the dataset contains missing values or irregular patterns.

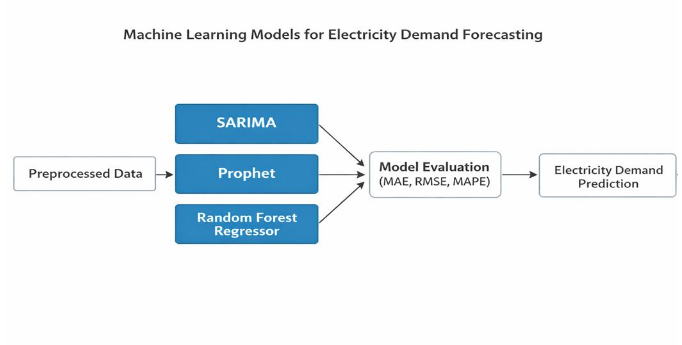


Figure 3: Machine Learning Model

C. *Random Forest Regressor*

Random Forest is an ensemble machine learning algorithm that builds multiple decision trees and combines their predictions to prod uce accurate results. It is capable of capturing nonlinear relationships between input features and electricity demand.
Model Evaluation
The performance of these models is evaluated using the following metrics:

- MAE (Mean Absolute Error)
- RMSE (Root Mean Square Error)
- MAPE (Mean Absolute Percentage Error)

The model with the lowest error values is selected as the best forecasting model for electricity demand prediction.

VI. EXPERIMENTAL RESULTS

The proposed electricity demand forecasting system was evaluated using historical electricity consumption data. The dataset was div ided into training and testing sets to analyze the performance of different forecasting models. Three models, namely SARIMA, Prop het, and Random Forest Regressor, were implemented and compared.
To measure prediction accuracy, standard evaluation metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error
(MAPE) were used. These metrics help determine how closely the predicted electricity demand values match the actual values.
The experimental results show that the Random Forest model achieved better forecasting accuracy compared to the SARIMA and Pr ophet models. This is because Random Forest can capture complex nonlinear relationships between electricity demand and time-bas ed features.
The results demonstrate that machine learning techniques can significantly improve electricity demand forecasting performance and provide reliable predictions for grid planning and energy management.

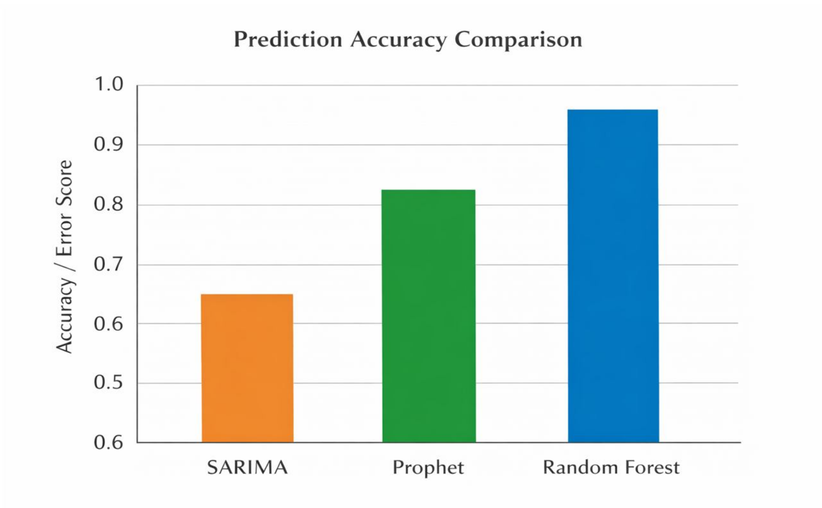


Figure 4: Prediction Accuracy Graph

VII. FUTURE SCOPE

The proposed electricity demand forecasting system can be further improved by integrating additional data sources such as weather conditions, economic indicators, and population growth, which can influence electricity consumption patterns. Future work may als o involve implementing advanced deep learning models such as Long Short-Term Memory
(LSTM) and GRU networks to capture long-term temporal dependencies in electricity demand data.
In addition, the system can be extended to support real-time electricity demand forecasting using live data streams from smart meter s and IoT devices. Deploying the forecasting system on cloud platforms can enable large-scale data processing and improve accessib ility for power utilities. Furthermore, integrating the forecasting framework with smart grid infrastructure can assist policymakers an d energy planners in making better decisions regarding power generation, renewable energy integration, and grid stability.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

VIII. DATASET DESCRIPTION

The dataset used in this study consists of historical electricity demand data collected over a specific time period. The dataset includes time-series records that represent electricity consumption at different timestamps. Each record contains information such as date, time, and electricity demand values, which are used to analyze consumption patterns and forecast future demand.

Before applying forecasting models, the dataset undergoes data preprocessing to remove missing values, convert datetime formats, and ensure data consistency. Additional time-based features such as year, month, day, and hour are extracted from the timestamp to capture seasonal and temporal variations in electricity consumption.

Furthermore, lag features are generated to represent previous electricity demand values, which help machine learning models learn patterns from past observations. The processed dataset is then divided into training and testing sets to evaluate the performance of forecasting models

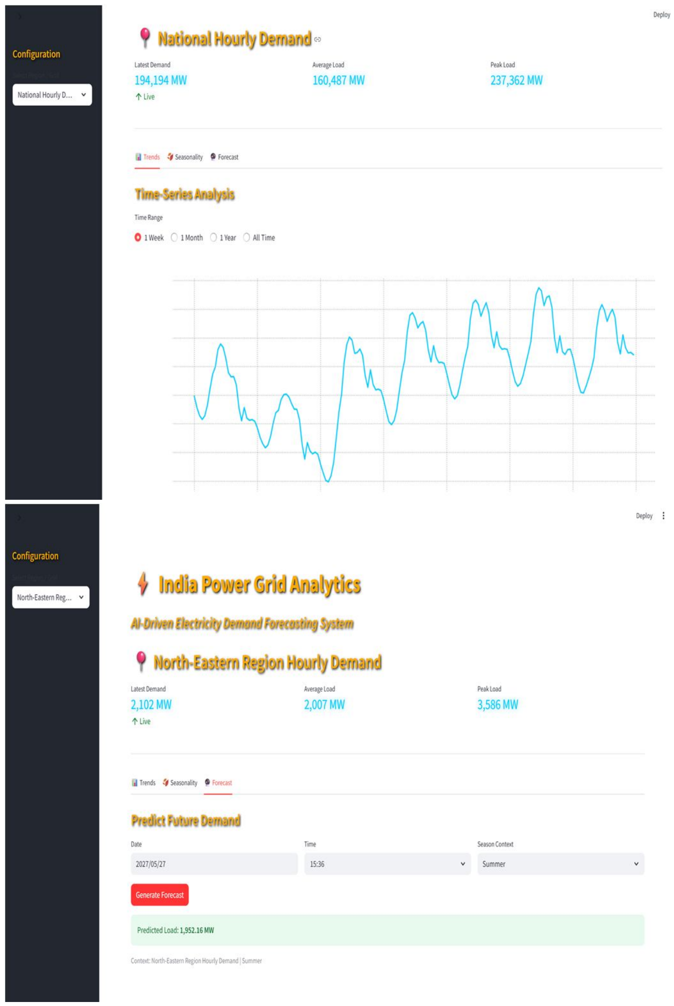


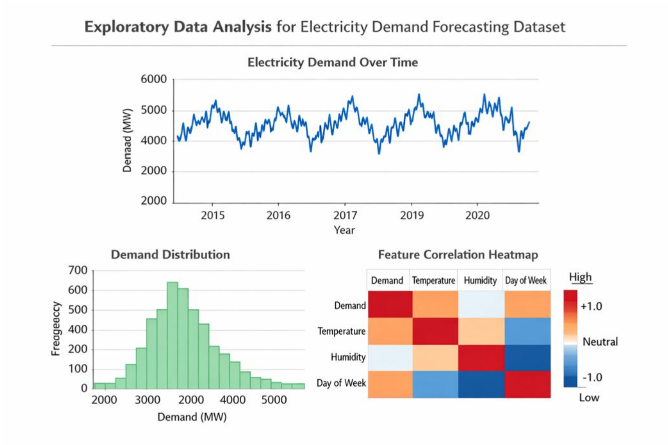
Figure 5: Output samples

IX. EXPLORATORY DATA ANALYSIS

(EDA) is performed to understand the characteristics and patterns present in the electricity demand dataset. In this stage, the dataset is analyzed using statistical summaries and visualization techniques to identify trends, seasonal patterns, and anomalies in electricity consumption.

Initially, the dataset is inspected to check for **missing values, duplicate records, and data inconsistencies**. After cleaning the dataset, summary statistics such as **mean, median, standard deviation, minimum, and maximum values** are calculated to understand the distribution of electricity demand.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com



Visual analysis is performed using graphs such as **line plots, histograms, and correlation heatmaps** to observe demand variations over time. Time-series plots help identify **daily, weekly, and seasonal demand patterns**, while correlation analysis helps determin e the relationship between different features and electricity demand.

X. ALGORITHM DETAILS

A. SARIMA Algorithm

SARIMA

(Seasonal AutoRegressive Integrated Moving Average) is used for time-series forecasting where seasonal patterns exist in the data.

Steps of SARIMA Algorithm

1) Collect historical electricity demand dataset.
2) Perform data preprocessing and convert the data into time-series format.
3) Identify seasonal and non-seasonal components using ACF and PACF plots.
4) Select SARIMA parameters p, d, q and seasonal parameters P, D, Q, s.
5) Train the SARIMA model using the training dataset.
6) Generate electricity demand forecasts.
7) Evaluate model performance using MAE, RMSE, and MAPE.

SARIMA Formula

$$SARIMA(p, d, q)(P, D, Q, s)$$

B. Prophet Algorithm

Prophet is a time-series forecasting model designed to handle trend and seasonal patterns in data.

Steps of Prophet Algorithm

1) Prepare the dataset with two columns: ds (date) and y (demand value).
2) Handle missing values and ensure correct datetime format.
3) Initialize the Prophet forecasting model.
4) Fit the model using historical electricity demand data.
5) Generate future time points for forecasting.
6) Predict electricity demand values for future timestamps.
7) Evaluate prediction accuracy using MAE, RMSE, and MAPE.

Prophet Model Formula

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

Where:

- $g(t)$ = trend
- $s(t)$ = seasonality
- $h(t)$ = holiday effects
- ϵ_t = error term

C. Random Forest Regressor Algorithm

Random Forest is an ensemble machine learning algorithm that combines multiple decision trees to improve prediction accuracy.

Steps of Random Forest Algorithm

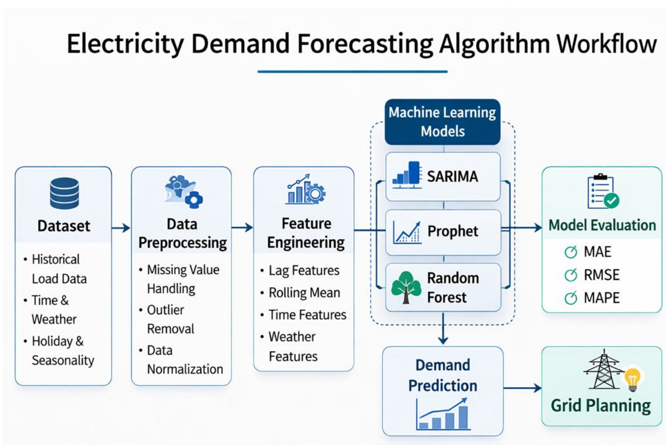
1) Import the processed electricity demand dataset.
2) Split the dataset into training and testing sets.
3) Select input features such as time-based variables and lag features.
4) Initialize the Random Forest Regressor model.
5) Train multiple decision trees using bootstrap sampling.
6) Each tree generates a prediction for electricity demand.
7) Combine predictions by averaging outputs from all trees.
8) Evaluate model performance using MAE, RMSE, and MAPE.

Random Forest Prediction Formula

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N} T_i\,(x)$$

Where

- N = number of trees
- $T_i(x)$ = prediction from each decision tree



XI. SYSTEM TESTING

System testing is performed to verify that the electricity demand forecasting system functions correctly and meets the specified requ irements. It ensures that all modules of the system such as data preprocessing, feature engineering, model training, and prediction ge neration work together without errors.

During system testing, different test cases were executed to validate each module of the system. The testing process includes checki ng whether the dataset loads correctly, verifying the presence of required columns, handling missing values, generating features, trai ning forecasting models, and evaluating prediction accuracy.

The forecasting models such as SARIMA, Prophet, and Random Forest were tested using historical electricity demand data. The res ults confirmed that the models were successfully trained and generated accurate predictionsEvaluation metrics such as Mean Absolu te Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error

(MAPE) were used to validate the prediction performance. All the modules passed the testing phase successfully, indicating that the system operates correctly and produces reliable electricity demand forecasts.

XII. ADVANTAGES OF THE SYSTEM

1) Accurate Demand Prediction:
   The system uses machine learning models such as SARIMA, Prophet, and Random Forest to provide more accurate electricity d emand forecasts.

2) Efficient Grid Planning
Forecasted electricity demand helps energy planners make better decisions for power generation, distribution, and grid infrastru cture.
3) Handles Large Datasets
Machine learning algorithms can process large volumes of historical electricity consumption data efficiently.
4) Captures Seasonal Patterns
Time-series models used in the system can identify seasonal and temporal patterns in electricity demand.
5) Improved Energy Management
Accurate forecasting helps reduce power shortages and improves overall energy management.
6) Supports Smart Grid Development
The system can be integrated with modern smart grid technologies for better monitoring and electricity distribution.
7) Scalable System
The forecasting system can be extended to include real-time data and additional features such as weather conditions.

XIII. LIMITATIONS

Although the proposed system performs effectively, certain limitations exist.
• Forecast accuracy depends on the quality and availability of historical data.
• The model may not capture sudden changes in electricity demand caused by unexpected events.
• Limited external factors such as weather or economic conditions may affect prediction performance.
• The system currently works mainly with historical datasets and not real-time data.

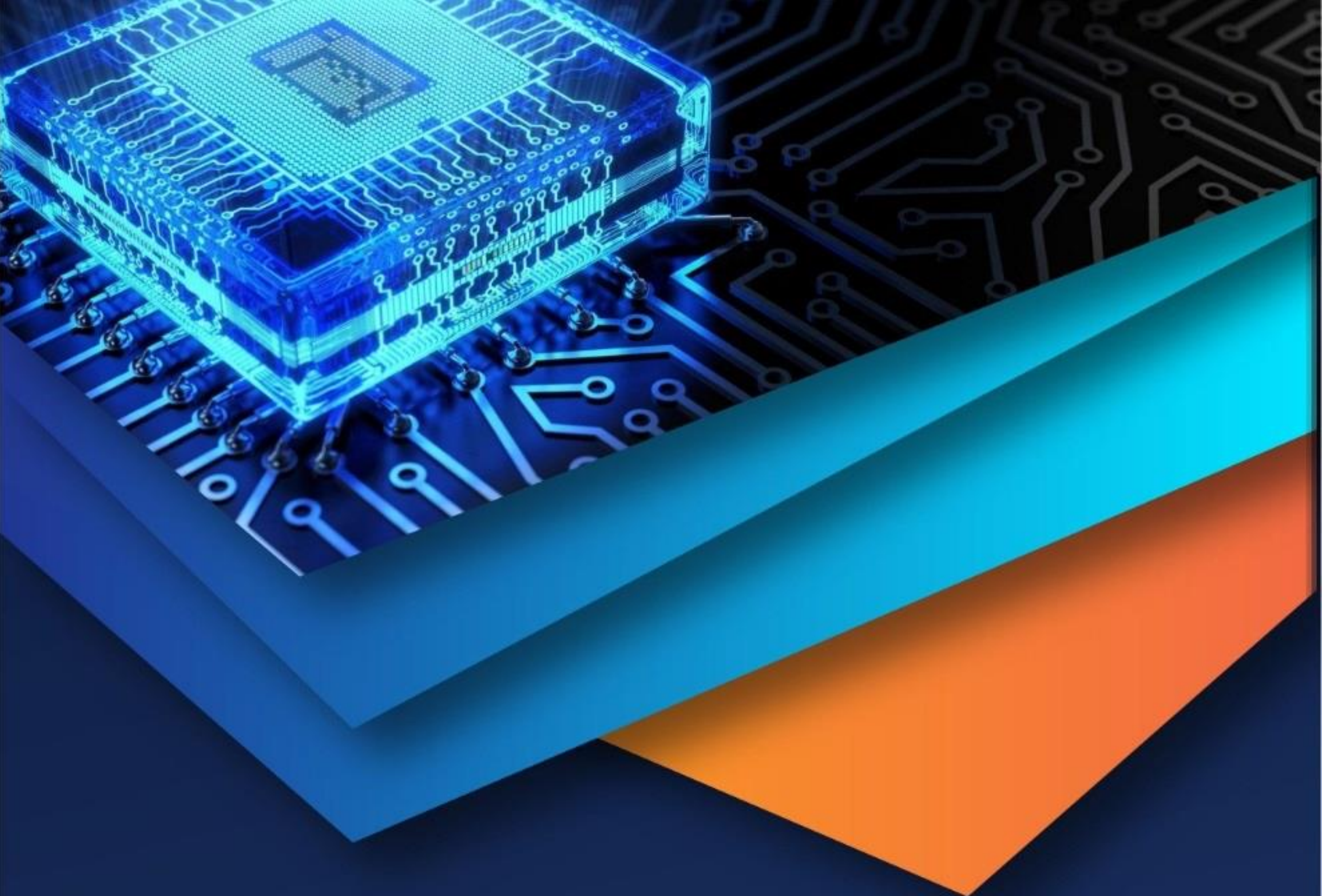
XIV. CONCLUSION

This study presented a machine learning based approach for electricity demand forecasting and grid planning for India. The system analyzes historical electricity consumption data and applies forecasting models such as SARIMA, Prophet, and Random Forest to pr edict future electricity demand.

The experimental results show that the Random Forest model provides better prediction accuracy compared to traditional statistical models. The proposed system can help power utilities and energy planners improve grid stability, energy management, and infrastru cture planning.

REFERENCES

[1] H. S. Hippert, C. E. Pedreira, and R. C. Souza,
Neural Networks for Short-Term Load Forecasting, IEEE Transactions on Power Systems, 2001.
https://ieeexplore.ieee.org/abstract/document/910780
[2] S. J. Taylor and B. Letham,
Forecasting at Scale, The American Statistician, 2018.
https://www.tandfonline.com/doi/abs/10.1080/00031305.2017.1380080
[3] T. Hong and S. Fan,
Probabilistic Electric Load Forecasting: A Tutorial Review, International Journal of Forecasting, 2016.
https://www.sciencedirect.com/science/article/abs/pii/S0169207015001508
[4] L. Breiman,
Random Forests, Machine Learning, 2001.
https://link.springer.com/article/10.1023/a:1010933404324
[5] R. J. Hyndman and G. Athanasopoulos,
Forecasting: Principles and Practice, 2018.
https://books.google.co.in/books?hl=en&lr=&id=_bBhDwAAQBAJ&oi=fnd&pg=PA7&dq=Forecasting+Principles+and+Practice+Hyndman&ots=TjlYwlVLIK&sig=POfbcROiBfhI4_4pwnbJQSCtzug&redir_esc=y - v=onepage&q=Forecasting%20Principles%20and%20Practice%20Hyndman&f=false
[6] G. Dudek,
Short-Term Load Forecasting Using Random Forests, 2015.
https://link.springer.com/chapter/10.1007/978-3-319-11310-4_71



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY