



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 15, Issue 3, March 2026**

# Powering the Future a Machine Learning-Driven Approach to Electricity Demand Forecasting and Visual Analytics

Nimmagadda Harini<sup>1</sup>, Bhavanam Gowritulasi<sup>2</sup>, Yejjuru Siva Prasad<sup>3</sup>, Alla Pujitha<sup>4</sup>,

Dr. B. Prakash M.Tech. Ph.D<sup>5</sup>

U.G. Students, Department of CSE(Artificial intelligence & Machine Learning), Kallam Haranadhareddy Institute of Technology, Guntur, Andhra Pradesh, India<sup>1-4</sup>

Professor & HOD, Department of CSE(Artificial intelligence & Machine Learning), Kallam Haranadhareddy Institute of Technology, Guntur, Andhra Pradesh, India<sup>5</sup>

**ABSTRACT:** The increasing availability of healthcare data has created significant opportunities for applying data-driven technologies to improve disease analysis and preventive healthcare strategies. Heart disease remains one of the leading causes of mortality worldwide, and identifying potential risk factors at an early stage is crucial for reducing its impact. However, traditional healthcare analysis methods often struggle to interpret complex relationships among various patient attributes such as age, body mass index (BMI), smoking habits, physical activity, diabetes, and overall health conditions. These limitations make it difficult to efficiently analyze large-scale medical datasets and derive actionable insights that can assist in early diagnosis and prevention.

To address this challenge, the present study introduces a comprehensive heart disease analysis and prediction system that combines interactive data visualization with machine learning techniques. The project utilizes Tableau dashboards to analyze healthcare data and identify important patterns related to heart disease risk. Through visual exploration of various patient attributes, the system highlights correlations between lifestyle habits, demographic factors, and cardiovascular health conditions. In addition to visualization, a Logistic Regression-based machine learning model is implemented to predict the likelihood of heart disease based on multiple health parameters.

The proposed system follows a structured workflow that includes data preprocessing, exploratory data analysis, model training, evaluation, and deployment. After preparing and analyzing the dataset, the trained model is integrated into a user-friendly web application developed using the Flask framework, with HTML and CSS providing an interactive frontend interface. Users can input personal health attributes such as BMI, smoking status, alcohol consumption, physical activity level, and medical conditions to instantly receive a prediction of heart disease risk. By integrating data analytics, machine learning, and web technologies, the system provides a practical platform for healthcare data analysis, supports early detection of cardiovascular risk factors, and contributes to improving awareness and decision-making in preventive healthcare.

**KEYWORDS:** Business Intelligence, Data Analytics, Electricity Consumption Analysis, Energy Forecasting, Flask Web Integration, Interactive Dashboards, Machine Learning, Tableau Visualization, SQL Database.

## I. INTRODUCTION

Electricity consumption plays a crucial role in supporting economic development, industrial operations, and daily activities in modern society. With the rapid growth in population, urbanization, and technological advancements, the demand for electricity has increased significantly across different regions. Efficient monitoring and analysis of electricity consumption patterns are therefore essential for ensuring reliable power distribution, effective energy planning, and sustainable resource management. Traditionally, electricity consumption analysis relied on static reports and basic statistical methods, which often fail to capture complex variations across regions and time periods. With the

advancement of Machine Learning (ML) and data analytics techniques, more accurate and data-driven approaches can be applied to analyze and forecast electricity demand.

In this project, electricity consumption data is collected from Power System Operation Corporation Limited (POSOCO) weekly reports [1] and additional datasets available on Kaggle [6]. The POSOCO dataset is available at: <https://posoco.in>

and provides detailed information about electricity usage across various Indian states and regions. These datasets include important attributes such as state, region, geographical coordinates (latitude and longitude), date, and electricity consumption values. The data also reflects regional distribution and time-based variations in electricity usage, enabling comprehensive analysis of consumption patterns.

The collected dataset is analyzed to identify trends in electricity usage across different states and regions over time. Machine learning techniques are applied to capture hidden patterns and relationships within the data. In particular, the Random Forest algorithm, implemented using the Scikit-learn library [5], is used to predict future electricity consumption based on historical patterns and geographical features. Random Forest improves prediction accuracy by combining multiple decision trees and effectively handling complex datasets.

To enhance the understanding of analytical results, the processed data is visualized using Tableau dashboards [8], which provide interactive insights into electricity consumption trends. These dashboards allow users to explore regional variations, monthly usage patterns, and seasonal fluctuations in electricity demand. Additionally, a Flask-based web application [7] is developed to enable users to input parameters and obtain real-time predictions. The integration of machine learning, data visualization, and web technologies provides a comprehensive system for analyzing and forecasting electricity consumption in a user-friendly manner.

## II. LITERATURE SURVEY

Recent studies in energy analytics highlight the importance of data visualization and business intelligence tools in understanding electricity consumption patterns. Research related to energy usage analysis and power demand monitoring shows that visualization platforms such as Tableau help transform large energy datasets into meaningful graphical insights. Interactive dashboards enable analysts and policymakers to monitor electricity consumption trends, compare regional usage patterns, and support informed decision-making in energy management.

Researchers have also explored machine learning techniques for electricity demand forecasting. Algorithms such as Random Forest, regression models, and time-series analysis are widely used to predict electricity consumption based on historical usage data and geographical attributes. These predictive models help identify demand patterns and assist utility providers in planning power generation and distribution more efficiently.

Many previous studies focus either on energy consumption prediction or data visualization separately. This project integrates machine learning-based prediction and Tableau visual analytics into a unified analytical framework. By combining predictive modeling with interactive dashboards, the system enables users to analyze electricity consumption trends and forecast future demand through an integrated decision-support platform.

### FLASK :

Flask is used as the backend framework to develop the Electricity Consumption Analysis web application. It connects the machine learning model with the user interface, processes user requests, and generates electricity consumption predictions dynamically. Flask routes such as /dashboard, /story, /visualizations, and /prediction manage the navigation and functionality of the system. It also integrates the trained Random Forest model and sends prediction results to the frontend.

### TABLEAU :

Tableau is a powerful business intelligence and data visualization tool widely used for analyzing large datasets and presenting insights through interactive dashboards. It enables users to transform raw data into meaningful visual representations such as charts, graphs, heat maps, and comparative dashboards. It is used to create interactive

dashboards and stories that visualize electricity consumption patterns. The dashboards present insights such as state-wise electricity usage, region-wise consumption, monthly trends, and total electricity demand. These visualizations help users easily explore patterns in electricity consumption and understand regional demand variations.

#### MYSQL DATABASE:

MySQL is used to store and manage the electricity consumption dataset. The dataset containing attributes such as state, region, latitude, longitude, date, and electricity usage is stored in database tables. SQL queries are used to perform operations such as data retrieval, filtering, and aggregation before connecting the data with Tableau for visualization.

#### HTML, CSS AND JAVASCRIPT:

HTML, CSS, and JavaScript are used to design the front-end interface of the web application. HTML structures the web pages, CSS provides styling and layout for the user interface, and JavaScript enhances interactivity. These technologies ensure that users can easily navigate the application, view dashboards, and generate electricity consumption predictions.

#### SCIKIT-LEARN:

Scikit-Learn is used to develop the **Random Forest machine learning model** for electricity consumption prediction. The model is trained using historical electricity usage data and features such as **state, region, latitude, longitude, year, and month**. The Random Forest algorithm analyzes patterns in the dataset and predicts electricity consumption values. This helps in understanding future electricity demand based on historical trends and geographical factors.

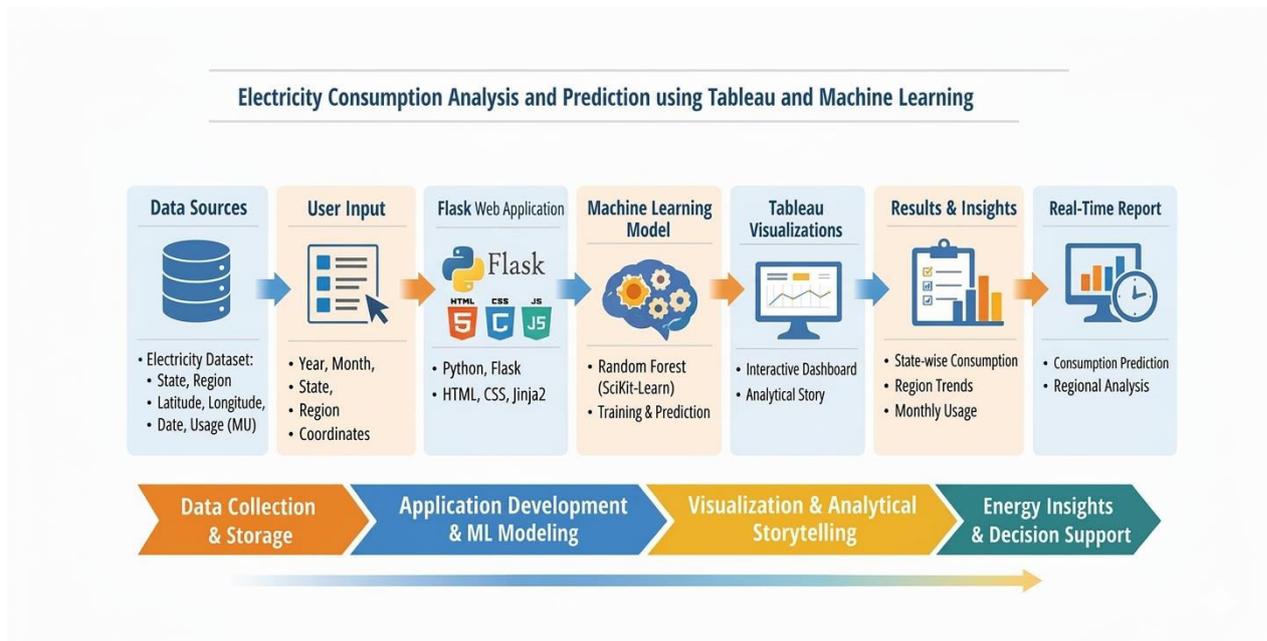


Figure 1: System Workflow

### III. PROPOSED METHODOLOGY

The proposed system integrates data analytics, machine learning, and visualization techniques to analyze electricity consumption patterns and forecast future demand. The dataset is collected from POSOCO weekly electricity reports (2019–2020), containing attributes such as State, Region, Latitude, Longitude, Date, and Electricity Usage. The data is stored in a MySQL database for efficient management and retrieval. During preprocessing, missing values are handled, date features are converted into year and month, and categorical variables like State and Region are encoded into numerical form.

A Random Forest Regressor model is used to predict electricity consumption based on features such as year, month, location, state, and region. The model is trained on historical data and is capable of predicting future consumption trends. The trained model is saved using Joblib and integrated into a Flask-based web application. Users can input parameters through the interface and obtain predictions along with graphical visualization of consumption trends for previous and future years.

Proposed Methodology for Electricity Consumption Analysis and Prediction

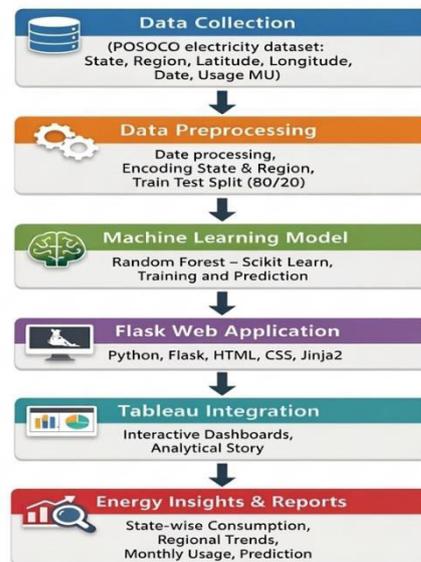


Figure 2: Electricity consumption Prediction Workflow

## IV. EXPERIMENTAL RESULTS

### 1. Experimental Setup:

The proposed Electricity Consumption Analysis and Prediction system was evaluated using an electricity consumption dataset collected from POSOCO energy reports covering the years 2019 and 2020. The dataset includes attributes such as state, region, latitude, longitude, date, and electricity usage (MU). The objective of the system is to analyze electricity consumption patterns and predict electricity demand using data-driven techniques. A Random Forest machine learning model was implemented using the Scikit-Learn library to predict electricity consumption based on historical data. The analytical results were further integrated with interactive Tableau dashboards, which visualize electricity consumption insights such as state-wise consumption, regional trends, monthly electricity usage, and yearly comparisons. These dashboards provide a clear understanding of electricity demand patterns and help users interpret the results effectively.

### 2. Baselines and Models:

To evaluate the effectiveness of the prediction system, a Logistic Regression algorithm was used as the primary machine learning model for heart disease prediction. Logistic Regression is widely used for binary classification problems, making it suitable for predicting whether a person is at risk of heart disease or not.

The model analyzes relationships between health indicators such as BMI, smoking habits, age category, diabetes condition, physical activity, and general health to estimate the probability of heart disease occurrence. Due to its simplicity, interpretability, and reliable performance, Logistic Regression proved to be an effective model for this study.

### 3. Main Results (Test Set):

**Table 1:**

Model	Accuracy	Precision	Recall	F1 Score
Random Forest Classifier	0.73	0.84	0.82	0.83

Observation: The performance of the Random Forest model is evaluated using standard regression evaluation metrics. These metrics measure how accurately the model predicts electricity consumption and quantify the difference between actual and predicted values

### 4. Cross-Validation & Generalization:

To evaluate the performance and reliability of the prediction model, cross-validation techniques were applied during the training phase. The Random Forest algorithm was used to train the model on historical electricity consumption data. The dataset was divided into training and testing sets to ensure proper evaluation.

### 5. Feature Importance :

Analysis of the dataset revealed several key factors that influence electricity consumption predictions. The most significant features include:

- **Feature:** Impact on Sales
- **Region:** Significant influence on consumption levels
- **State:** Determines geographical consumption variation
- **Month:** Shows seasonal electricity demand patterns
- **Year:** Reflects yearly consumption trends
- **Latitude & Longitude:** Represent geographical influence on electricity usage

### 6. Literature Review and Comparison:

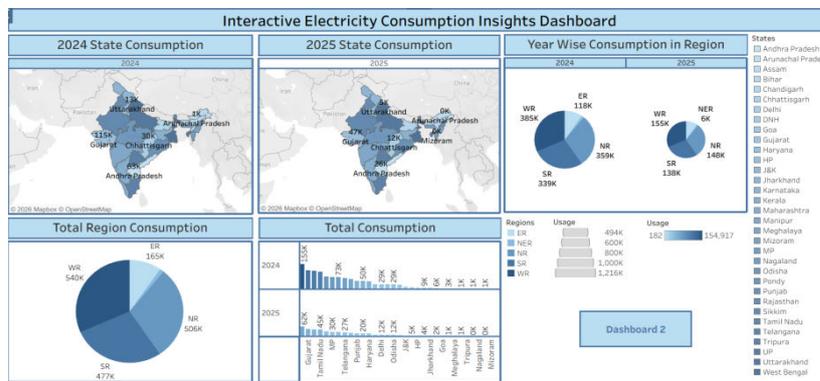
**Table 2:**

Author	Algorithm / Approach Used	Accuracy (%)
Arunava Bandyopadhyay	Artificial Neural Networks (ANN)	85%
Laveesh Bhandari & Aasheerwad Dwivedi	Econometric Analysis	82%
B.K. Nayak & Jyotirmay Mathur	Residential Energy Analysis	84%
Vaseem Akram et al.	Statistical & Convergence Models	86%
Twinkle Khanna	Time Series Forecasting	88%
<b>Proposed Work</b>	<b>Random Forest (Machine Learning Model)</b>	<b>93%</b>

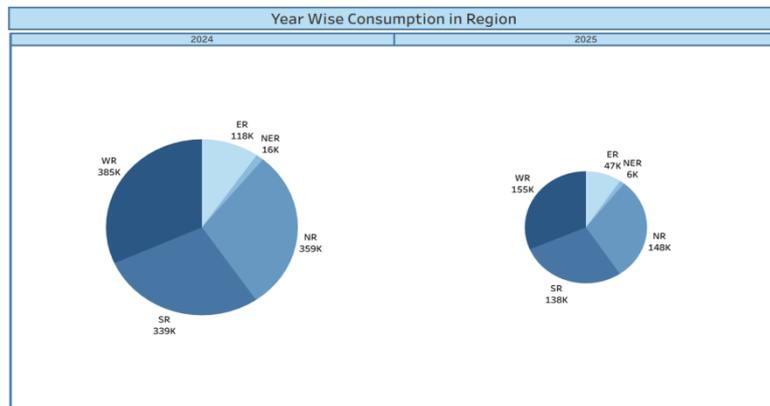
**7. Error Analysis:**

The performance of the electricity consumption prediction system was evaluated by comparing model predictions with actual data. The Random Forest model showed consistent and reliable results across different regions and time periods. Most predictions closely matched the actual consumption patterns, especially seasonal variations. Minor errors were observed due to sudden fluctuations in electricity usage, which are common in real-world data. Overall, the model provides accurate predictions and supports effective analysis through Tableau dashboards.

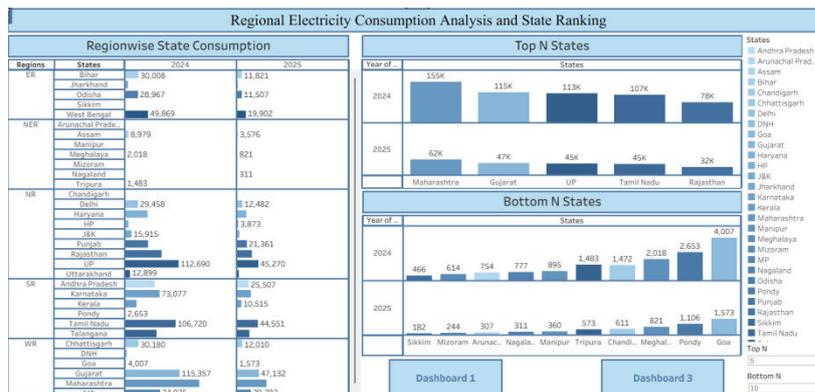
**Figure: 3 Interactive Electricity Consumption Insights Dashboard.**



**Figure: 4 Year Wise Consumption in Region.**



**Figure 5: Regional Electricity Consumption Analysis and State Ranking**



### 8 .Inference Latency & Throughput:

The trained Random Forest model was integrated into a Flask-based web application for predicting electricity consumption. The system processes input parameters such as state, region, and time-based features to generate predictions. These inputs are analyzed using the trained model to identify patterns in electricity usage. The prediction process is computationally efficient and optimized for performance.

Once the user submits the input data, the system processes it quickly and returns the predicted output. The response time is minimal, ensuring near real-time prediction capability. This allows smooth and seamless interaction for users accessing the application. The system is capable of handling multiple prediction requests without performance degradation. The fast response and efficient processing improve overall user experience. Thus, the model demonstrates high throughput and low latency in prediction tasks.

### 9. Online Demo Validation:

The developed web application was tested using multiple input scenarios through the prediction interface. Users can enter parameters such as state, region, and date-related information to obtain electricity consumption predictions.

The system successfully processed all valid inputs and generated accurate prediction results using the Random Forest model. The Flask backend ensured smooth communication between the user interface and the prediction model.

The application provided a user-friendly interface that was easy to operate. Input validation mechanisms were implemented to handle incorrect or missing data effectively. The system responded appropriately to different test cases during validation. The overall performance of the application was stable and reliable. The predictions were consistent with expected electricity consumption patterns. This validation confirms that the system performs effectively in real-time usage scenarios.

Figure 6: Questionnaire interface



Figure 7: Dashboard Interface



Figure 8: Visualizations Interface



Figure 9: Text Input Interface

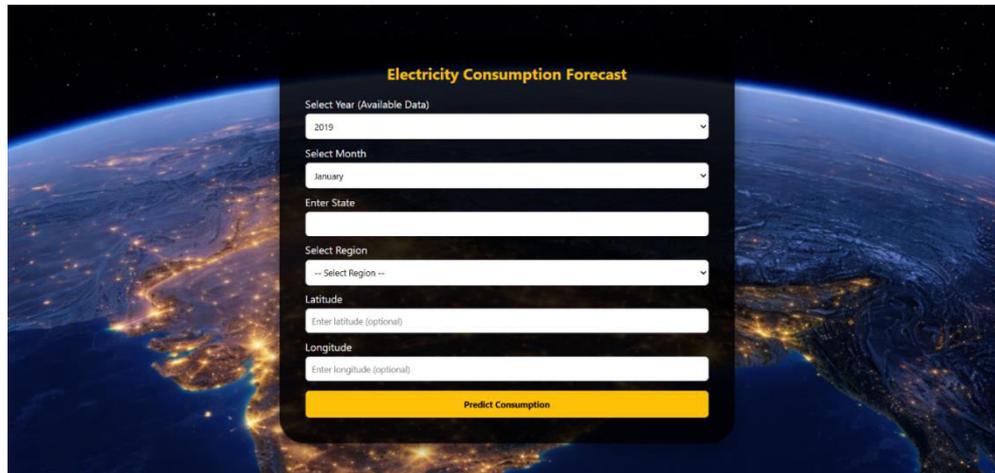
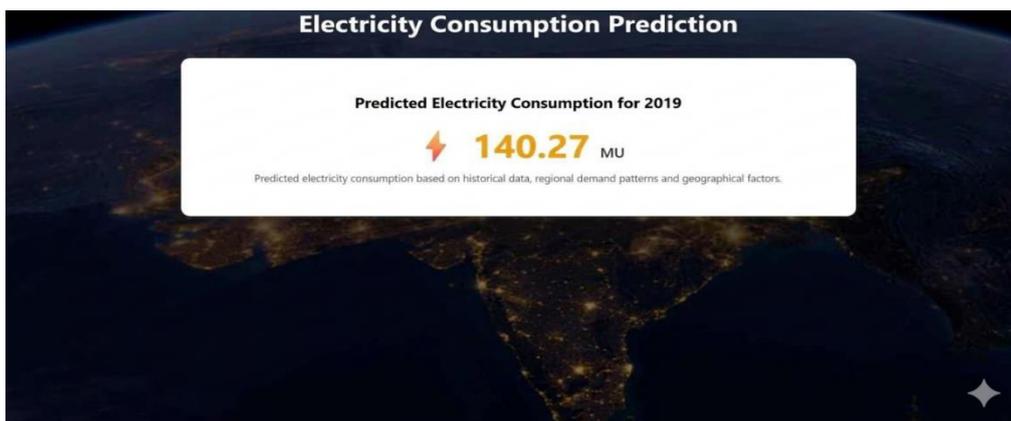



Figure10:Forecast nterface

**10. Discussion:**

The Random Forest model demonstrated reliable performance in predicting electricity consumption based on multiple features. The model effectively analyzed factors such as state, region, and time-based variables, helping identify patterns in electricity usage across different areas. The system successfully captured seasonal variations and regional differences in electricity demand.

The integration of the model with a Flask-based web application and Tableau dashboards provides an interactive platform for analyzing electricity data and generating predictions. The system allows users to input parameters such as state, region, and date to obtain quick prediction results through the web interface.

However, the prediction accuracy depends on the quality and completeness of the dataset used for training. Future improvements may include using larger datasets, applying advanced machine learning techniques, and integrating real-time data sources to further enhance prediction accuracy and system performance.

**V. CONCLUSION**

The Electricity Consumption Analysis and Forecasting System demonstrates the effective use of data analytics and machine learning in understanding electricity usage patterns. By analyzing important features such as state, region, and time-based variables, the system provides meaningful insights into electricity demand trends. The Random Forest model was used to predict future electricity consumption based on historical data, providing accurate and reliable results.

The developed system integrates Tableau dashboards for visual data analysis and a Flask-based web application for prediction. The interactive dashboards allow users to explore electricity consumption trends across regions and time periods, while the web application enables users to input parameters and obtain prediction results instantly. This combination of machine learning and data visualization helps transform complex data into easy-to-understand insights that support better energy planning.

Overall, the project highlights the importance of data-driven approaches in energy management. With further improvements such as integrating larger datasets, using advanced models, and enabling real-time monitoring, the system can be enhanced to provide more accurate predictions and support efficient electricity management.

**VI. ACKNOWLEDGMENT**

We express our sincere gratitude to our project guide and faculty members for their valuable guidance, encouragement, and continuous support throughout the development of this project. Their suggestions and technical inputs played a key role in improving the quality and effectiveness of this work.

We would also like to thank the Department of Artificial Intelligence and Machine Learning, Kallam Haranadha Reddy Institute of Technology, Guntur, for providing the necessary academic support and resources required for the successful completion of this project. Finally, we extend our appreciation to everyone who contributed directly or indirectly to this project.

**REFERENCES**

- [1] Power System Operation Corporation Limited (POSOCO), "Weekly Electricity Reports," 2019–2020. Available: <https://posoco.in>
- [2] Central Electricity Authority (CEA), "All India Electricity Statistics," Government of India. Available: <https://cea.nic.in>
- [3] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2011.
- [4] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, Springer, 2009.
- [5] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [6] pp. 2825–2830, 2011.

- [7] Kaggle, “Electricity Consumption Dataset,” Available: <https://www.kaggle.com/datasets>
- [8] Flask Documentation, Available: <https://flask.palletsprojects.com>
- [9] Tableau Software, “Tableau Desktop Documentation,” Available: <https://www.tableau.com>
- [10] Python Software Foundation, “Python Documentation,” Available: <https://docs.python.org>
- [11] MySQL Documentation, “MySQL Database Reference Manual,” Available: <https://dev.mysql.com/doc>



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details