



---

## **AI-BASED ENERGY CONSUMPTION FORECASTING IN SMART GRIDS FOR FUTURE IOT-BASED ENERGY METERS**

J. Prashanthi<sup>1</sup>, Ch. Poojitha<sup>2</sup>, T. Maruthi Megansh<sup>2</sup>, K. shiva kumar<sup>2</sup>, Amol rathod<sup>2</sup>

<sup>1</sup>Assistant Professor, <sup>2</sup>UG Student, <sup>1,2</sup>Department of Computer Science and Engineering (Data Science)

<sup>1,2</sup>Sree Dattha Group of Institutions, Sheriguda, Ibrahimpatnam, 501510, Telangana.

---

Received: 09-07-2025

Accepted: 23-08-2025

Published: 30-08-2025

---

### **ABSTRACT**

The integration of smart grids with Internet of Things (IoT) technology has transformed energy management systems by enabling real-time monitoring and efficient control of power distribution. As IoT-based energy meters become increasingly prevalent, accurately forecasting energy consumption has become critical for optimizing load balancing, minimizing energy waste, and enhancing demand response mechanisms. However, predicting energy usage is challenging due to varying consumer behaviors, environmental fluctuations, and seasonal trends. Traditional forecasting methods such as ARIMA and moving averages are limited by their inability to handle non-linear patterns, adapt to large-scale data, or incorporate diverse features like weather data and time-based variables. These methods often struggle with noise, outliers, and dynamic consumption trends, resulting in poor prediction accuracy. Therefore, there is a growing need for intelligent, scalable, and adaptive systems that can analyze high-dimensional data generated by smart meters. Addressing these limitations, the proposed system introduces a machine learning-based web application that uses Random Forest Regressor and Support Vector Regression (SVR) to forecast energy consumption. The system covers the complete pipeline, including preprocessing, outlier handling, feature selection, exploratory data analysis, model training, and evaluation, all integrated into an interactive Flask-based web interface. It efficiently captures the non-linear dependencies between various factors such as temperature, humidity, time, and calendar events with energy usage patterns. The use of advanced machine learning models improves forecasting accuracy, enabling power providers and stakeholders to make data-driven decisions, ensure grid stability, and support the development of sustainable energy practices in smart cities. By overcoming the limitations of traditional methods and effectively leveraging the richness of IoT-generated data, the proposed solution represents a significant step toward intelligent energy forecasting in next-generation smart grids.

### **1. INTRODUCTION**

The integration of machine learning-based energy consumption forecasting in smart grids marks a pivotal advancement in the energy sector, addressing the limitations of traditional forecasting methods that struggle with the dynamic, complex nature of modern energy systems. This research is motivated by the urgent need to enhance forecasting accuracy, which is critical for grid stability, cost reduction, efficient renewable energy integration, and consumer empowerment. Traditional systems often rely on statistical models that lack adaptability and fail to incorporate vital contextual data such as

weather or user behavior, resulting in suboptimal performance. By developing advanced machine learning models—leveraging regression algorithms, ensemble learning, and diverse datasets—this study aims to deliver robust, scalable, and data-driven solutions that improve operational efficiency and support sustainable energy practices. The applications of such models span across utilities, smart buildings, and renewable energy management, showcasing their transformative potential in optimizing energy distribution, promoting conservation, and enabling a more resilient and intelligent energy infrastructure.

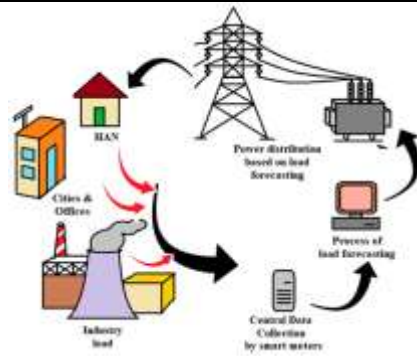


Figure 1: Application of energy consumption.

## 2. LITERATURE SURVEY

The invention of the Internet of Things (IoT) is one of the most significant technological advances of the 21st century. The IoT is a network of linked hardware, software, and physical nodes that enables data gathering and distribution. The exchanging of data amongst multiple infrastructures and devices is referred to as the “IoT” [1]. Without them, Internet access is impossible. The Internet Protocol monitors Internet devices. The Internet allows user-to-user online communication. The networked “things” aspire to provide everyday objects with the ability to share data and information on their own and at regular intervals. A structure that uses connected data, technology, and machinery is called an “intelligent building.” Analytics and automation for controlling essential services, including HVAC, lighting, heating for safety equipment, and air conditioning HVAC devices with intelligent controls [2]. It is highly beneficial for HVAC systems to utilize less energy if they have authorities implemented when energy is in high demand. The problematic parts of this duty include recognizing and locating defects and minimizing power consumption in vacant portions of the facility [3].

The integration of renewable energy sources, Smart Grid (SG) management, energy trading, power system flexibility and negative pricing, energy management, arbitrage and pricing, and SG financial transactions are some recommendations for enhancing the performance of SG block chain technology

and cryptocurrencies may lead to a flatter load profile and economic advantage. In addition, previous researchers [make changes to the block chain’s algorithms so that communities of online users may use Bitcoin to transact in marketplaces and energy systems using more secure methods [4]. The writers also use digital money in their energy infrastructure. A longer-term viewpoint may be just as helpful as a more immediate one when attempting to detect and mitigate possible hazards. The writers pay little attention to enduring challenges, well-known problems, or feasible remedies. Using block chain-based digital currencies, participants in the energy industry must solve security problems if they want to be effective and competitive.

Many more articles are available that cover a wide range of topics that impact all human cultures. For instance, the essay highlights the issue of global warming, which is leading to substantial changes in almost every part of the world. In this research, the electrical power utilized by ordinary home appliances is broken down using the Electrical Line Disaggregation (ELD) method. ELDs today depend on computer science techniques and Artificial Intelligence (AI). Optimized complete set empirical model decomposition and wave packet transformation, or OCEEMD-WPT, is also used. This idea was developed to show how the end user might perceive changes in power-line noise. Consequently, gathering vital information required for network operation is significantly more effective [5].

Using sensors or smart meters, a power grid might become an SG. These robust sensors transmit a lot of data. This helps understand network behavior and make assumptions. The vast data required to join and store thousands of IoT nodes makes this impossible [6]. Automatic Encoders (AE) approach encoding data entropy to represent previously compressed content with fewer data. These strategies use AI and deep learning. Until recently, data spectrum made compression

impossible. The suggested data compression method leverages AE models. Spectral windows improve compression and entropy.

IoT devices and technology may improve SG via real-time monitoring, new pricing methods, dynamic energy management, and self-healing. Intrusions are more significant in SG-converted grid components and services. Researchers studied attacker and defender payoffs using actual devices and honeypots. Both attacked and defended games contain uncertain NE and Bayesian NE matching conditions. The authors suggested increasing worst-case outcomes in non-equilibrium circumstances. If the defendant accepts the onslaught and gives up, he may submit. Simulations show that both games were balanced offensively and defensively. Defense recognizes and rewards aggressors. By interacting with a certain number of actual devices and honeypots, previous research looked at how an attacker and defense may cooperate to maximize their payoffs [7].

SGs use automation, sensors, and remote controls to increase comfort, security, and energy efficiency. IoT sensors monitor “smart” construction elements. This knowledge can enhance interiors. IoT-based “SG’s,” which monitor a building’s temperature, security, and maintenance, are made more accessible by smartphones and tablets. The IoT’s ability to link many sensors allows it to collect and analyze data in real time, leading to more innovative and user-friendly buildings [8]. For SGs, fire alarms are essential. A smart IoT fire alarm system is required to prevent property damage and save lives. Weka and J48 are used; previous research demonstrated energy-use patterns and behaviors. These were then categorized according to how much energy they used [9]. With machine learning and big data for the home, the HEMS-IoT smart energy management system lowers the home’s overall energy consumption while ensuring the comfort and security of its residents. The system relies heavily on machine learning and

large amounts of data to analyze and categorize how effectively energy is utilized, identify trends in human behavior, and maintain a high degree of comfort for building occupants [10].

### 3. PROPOSED METHODOLOGY

energy consumption forecasting is increasingly vital as global electricity demand surges due to urbanization, industrialization, and the proliferation of electric vehicles and digital devices, with the IEA reporting a 6% rise in global electricity usage in 2021 alone. Smart grids, empowered by IoT-based energy meters, offer real-time monitoring and management of energy flow, generating large-scale time-series data critical for data-driven decision-making. This forecasting is essential for optimizing resource allocation, preventing grid overloads, and implementing dynamic pricing and demand response strategies. The proposed system architecture introduces a novel hybrid AI-based model specifically designed for IoT-enabled smart grids, combining multi-source energy data with advanced preprocessing such as anomaly-resilient filters, dynamic feature scaling, and temporal EDA methods including trend decomposition and heatmap correlations. A stratified, time-aware train-test split preserves temporal consistency for real-world deployment, while a customized Random Forest Regressor—enhanced with adaptive ensemble tuning based on data volatility—outperforms traditional SVR in dynamic scenarios. This end-to-end forecasting solution is validated using real-time IoT feeds and evaluated with MAE, RMSE, and  $R^2$  metrics, offering an explainable, scalable, and application-specific advancement over prior methods.

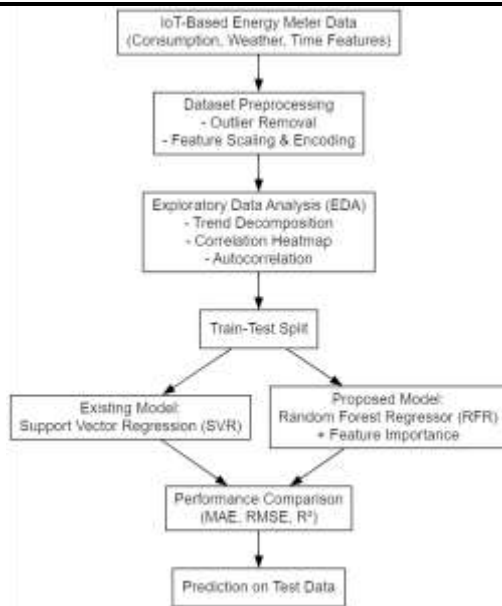


Figure 2: Proposed System Architecture.

### 3.2 Preprocessing

This application-specific preprocessing pipeline for energy consumption forecasting begins by parsing the “DateTime” column into a datetime object and decomposing it into Month, Day, Hour, Minute, and Year, then drops the original timestamp; next, it separates input features from the “TotalUsage” target and applies IQR-based Winsorization to cap extreme temperature and humidity values, preserving distribution integrity; following outlier treatment, it uses SelectKBest with an ANOVA F-test to retain the ten most predictive features before standardizing them via z-score normalization to ensure uniform scale; alongside these steps, an EDA phase computes correlations (visualized in a heatmap) and distribution plots with kernel density estimates to reveal variable relationships and inform any further transformations; finally, the cleaned and engineered dataset is partitioned—typically in an 80:20 or 70:30 stratified split—into training and test sets so that model performance (measured by MSE, MAE, or  $R^2$ ) can be reliably evaluated on unseen data, thereby bolstering robustness and generalization.

### 3.3 RFR (Random Forest Regressor)

Random Forest Regressor is a highly effective method for forecasting energy consumption in

smart grids due to its ability to handle complex, non-linear relationships and noisy IoT data without extensive preprocessing. This ensemble learning approach combines multiple decision trees, reducing overfitting and improving predictive accuracy, which is crucial for application-specific data where consumption patterns can vary widely across households or industries. RFR can naturally handle large feature sets, including categorical and continuous variables, making it suitable for integrating diverse data sources such as time, weather, and usage history. Its robustness and interpretability help energy providers build reliable forecasting models that adapt to the dynamic nature of smart grid environments.

#### Step 1: Building Multiple Decision Trees

The process begins by creating numerous decision trees from different random subsets of the training data and features. Each tree learns to predict energy consumption by recursively splitting the data based on feature values to minimize prediction error. This randomness ensures that each tree captures different aspects and variations in the data, improving generalization.

#### Step 2: Aggregating Predictions

Once all trees are trained, the Random Forest model aggregates their outputs by averaging the predictions from individual trees. This ensemble averaging reduces the risk of overfitting to noise or anomalies present in specific data subsets, producing a more stable and accurate forecast.

#### Step 3: Model Evaluation and Application

The aggregated prediction is then evaluated on test data using error metrics such as Mean Squared Error, Mean Absolute Error, and R-squared score to assess performance. Based on these metrics, energy providers can fine-tune the model’s parameters, such as the number of trees or maximum tree depth, to better fit the application context. The resulting model can then be deployed to generate reliable energy consumption forecasts, aiding in efficient grid management and planning.

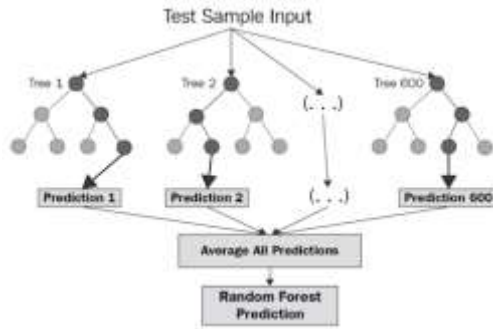


Figure 3: Proposed RFR Regression.

## 4.RESULTS

Figure 4 shows the homepage of the "Smart Grid Energy Consumption Forecasting" web application, accessible at 127.0.0.1:5000/. The page displays the application title in bold at the top, followed by a navigation bar with six buttons: "Dataset," "Preprocessing," "EDA," "Algorithms," "Performance Comparison," and "Prediction From Test." These buttons link to different functionalities of the application, allowing users to explore the dataset, preprocess data, perform exploratory data analysis, view trained models, compare model performance, and make predictions. The layout is minimalistic, with a white background and blue buttons, providing a clean interface for user interaction.



Figure 4: Home page

Dataset													
TemperatureF	Humidity	Hour_y	Minute_y	Day_y	Weekend	Holiday	TotalUsage	Month	Hour	Minute	Day	Year	
50.0	63.0	0	0	6	0	0	19.843233	1	0	0	1	2016	
50.0	63.0	1	0	6	0	0	19.843233	1	1	0	1	2016	
50.0	63.0	2	0	6	0	0	19.843233	1	2	0	1	2016	
50.0	63.0	3	0	6	0	0	19.843233	1	3	0	1	2016	
50.0	63.0	4	0	6	0	0	19.843233	1	4	0	1	2016	

Figure 5: Uploaded Dataset

Figure 5 displays the "Dataset" page, showing the first five rows of the dataset loaded from Dataset.csv. The table includes columns: TemperatureF, Humidity, Hour\_y, Minute\_y,

Day\_y, Weekend, Holiday, TotalUsage, Month, Hour, Minute, Day, and Year. The first row, for example, has values: TemperatureF of 50.0°F, Humidity of 63.0%, Hour\_y of 0, Minute\_y of 0, Day\_y of 6, Weekend of 0, Holiday of 0, TotalUsage of 19.843233, Month of 1, Hour of 0, Minute of 0, Day of 1, and Year of 2016. This page provides a snapshot of the dataset, with a "Back to Home" button for navigation.



Figure 6: Preprocessing

Figure 6 presents the "Preprocessing" page, which summarizes missing values and preprocessing steps for the dataset. It lists missing values for each column, all showing 0: TemperatureF (0), Humidity (0), Hour\_y (0), Minute\_y (0), Day\_y (0), Weekend (0), Holiday (0), TotalUsage (0), Month (0), Hour (0), Minute (0), Day (0), and Year (0). Below this, a note describes the preprocessing steps: parsing the DateTime column, extracting features (Month, Hour, Minute, Day, Year), handling outliers with Winsorizer, selecting features with SelectKBest, and standardizing with StandardScaler. A "Back to Home" button is provided.

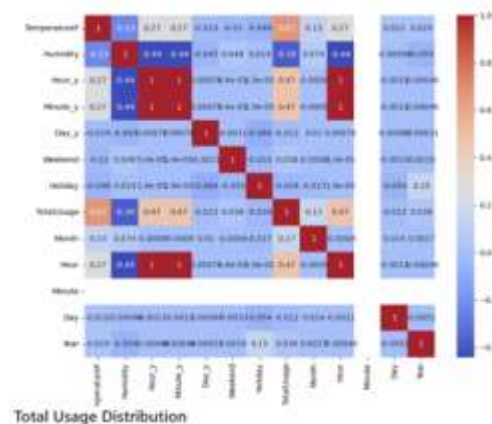


Figure 7: Correlation Heatmap

Figure 7 illustrates the "Exploratory Data Analysis" page, specifically the "Correlation Heatmap" section. The heatmap visualizes correlations between dataset features, with values ranging from -0.4 to 1.0. Notable correlations include: Hour\_y and Minute\_y at 1.0 (perfect correlation, indicating redundancy), Hour and Hour\_y at 1.0, Minute and Minute\_y at 1.0, Day and Day\_y at 1.0, and TotalUsage with Hour at 0.21 (moderate positive correlation). Negative correlations include TemperatureF with Humidity at -0.27. The color scale ranges from blue (negative) to red (positive), aiding in identifying relationships between features.

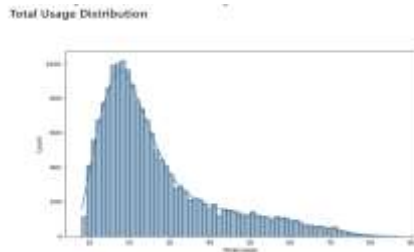


Figure 8: Total Usage

Figure 8 shows the "Total Usage Distribution" section of the EDA page, presenting a histogram with a kernel density estimate (KDE) for the Total Usage column. The x-axis represents TotalUsage values (ranging from 0 to 90), and the y-axis shows the count (up to 1000). The distribution is right-skewed, with a peak around 10–20 (count ~900), indicating that most TotalUsage values are low, with fewer instances of higher usage (e.g., a small spike near 80). This visualization highlights the skewed nature of energy consumption, useful for understanding the target variable's distribution.



Figure 9: Algorithms used

Figure 9 displays the "Algorithms" page, listing the trained models used for prediction.

Two models are shown: "RandomForest: RandomForestRegressor(random\_state=42)" and "SVR: SVR()". These models predict TotalUsage based on preprocessed features. The RandomForestRegressor is configured with a random\_state of 42 for reproducibility, while the SVR uses default parameters. A "Back to Home" button allows navigation back to the homepage. This page provides transparency on the machine learning models employed.



Figure 10: performance Comparison

Figure 10 presents the "Performance Comparison" page, showing a table comparing the two models' performance on the test set. The table includes three metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), and R<sup>2</sup> score. For RandomForest, the values are MSE of 13.4051, MAE of 2.7197, and R<sup>2</sup> of 0.9358. For SVR, the values are MSE of 19.6689, MAE of 3.2909, and R<sup>2</sup> of 0.9057. A "Back to Home" button is included. This comparison helps users evaluate model effectiveness.



Figure 11: Prediction From Test

Figure 11 shows the "Prediction From Test" page, where users can input values to predict TotalUsage. The form includes fields for all features: TemperatureF (0 to 100), Humidity (0 to 100), Hour\_y (0 to 23), Minute\_y (60 to 1380, multiple of 60), Day\_y (1 to 7), Weekend (0 or 1), Holiday (0 or 1), Month (1 to 12), Hour (0 to 23), Minute (0 to 59), Day

(1 to 31), and Year (2025 or earlier). A sample prediction is shown with a Predicted Total Usage of 16.5367, likely using the RandomForest model. A "Predict" button submits the form, and a "Back to Home" button is provided.

## 5. CONCLUSION

The "Smart Grid Energy Consumption Forecasting" web application effectively facilitates energy usage prediction by leveraging a dataset with environmental and temporal features to train Random Forest and SVR models, achieving high predictive accuracy with  $R^2$  scores of 0.9358 and 0.9057, respectively. The application provides a user-friendly interface for dataset exploration, preprocessing (including datetime parsing, outlier handling with Winsorizer, feature selection via Select-K-Best, and standardization), exploratory data analysis through visualizations like correlation heatmaps and usage distributions, and performance comparison, with Random Forest outperforming SVR across all metrics (MSE of 13.4051 vs. 19.6689, MAE of 2.7197 vs. 3.2909). The prediction functionality allows users to input feature values and obtain accurate total usage forecasts, demonstrating the application's practical utility for smart grid management.

## REFERENCES

- [1]. Zúquete, A.; Gomes, H.; Amaral, J.; Oliveira, C. Security-Oriented Architecture for Managing IoT Deployments. *Symmetry* 2019, 11, 1315.
- [2]. Nappi, I.; de Campos Ribeiro, G. Internet of Things technology applications in the workplace environment: A critical review. *J. Corp. Real Estate* 2020, 22, 71–90.
- [3]. Dos Santos, D.R.; Dagrada, M.; Costante, E. Leveraging operational technology and the Internet of things to attack smart buildings. *J. Comput. Virol. Hacking Tech.* 2021, 17, 1–20.
- [4]. Ghorbanian, M.; Dolatabadi, S.H.; Siano, P.; Kouveliotis-Lysikatos, I.; Hatziargyriou, N.D. Methods for flexible management of blockchain-based cryptocurrencies in electricity markets and smart grids. *IEEE Trans. Smart Grid* 2020, 11, 4227–4235.
- [5]. Chui, K.T.; Gupta, B.B.; Liu, R.W.; Vasant, P. Handling data heterogeneity in electricity load disaggregation via complete optimized ensemble empirical mode decomposition and wavelet packet transform. *Sensors* 2021, 21, 3133.
- [6]. Lee, J.; Yoon, S.; Hwang, E. Frequency Selective Auto-Encoder for Smart Meter Data Compression. *Sensors* 2021, 21, 1521.
- [7]. Diamantoulakis, P.; Dalamagkas, C.; Radoglou-Grammatikis, P.; Sarigiannidis, P.; Karagiannidis, G. Game-theoretic honeypot deployment in smart Grid. *Sensors* 2020, 20, 4199.
- [8]. Khajenasiri, I.; Estebasari, A.; Verhelst, M.; Gielen, G. A review on Internet of Things solutions for intelligent energy control in buildings for smart city applications. *Energy Procedia* 2017, 111, 770–779.
- [9]. Machorro-Cano, I.; Alor-Hernández, G.; Paredes-Valverde, M.A.; Rodríguez-Mazahua, L.; Sánchez-Cervantes, J.L.; Olmedo-Aguirre, J.O. HEMS-IoT: A big data and machine learning-based smart home system for energy saving. *Energies* 2020, 13, 1097.
- [10]. Mohanta, B.K.; Jena, D.; Satapathy, U.; Patnaik, S. Survey on IoT security: Challenges and solution using machine learning, artificial intelligence, and blockchain technology. *Internet Things* 2020, 11, 100227.