ELECTRICITY USAGE AND BILL PREDICTION SYSTEM BASED ON HOUSEHOLD ATTRIBUTES AND APPLIANCE CONSUMPTION

Mrs. Divya M,

Department of CSE

Rajalakshmi Engineering College

chennai,India

divya.m@rajalakshmi.edu.in

Neela A(220701184)

Department of CSE

Rajalakshmi Engineering College

Chennai,India

220701184@rajalakshmi.edu.in

Abstract: In an era of rising energy demands and increasing utility costs, the ability to predict electricity consumption and billing has become a critical aspect of sustainable living and energy management. This paper presents a machine learning-based predictive system designed to estimate future electricity usage and bill amounts by analyzing household-specific features. These include house size, number of residents, historical electricity bills, weather conditions, and most importantly, the frequency and intensity of heavy appliance usage. By training regression models on historical data and real-world usage patterns, the system delivers accurate and customized forecasts that help users anticipate and control their energy expenses.

The proposed system stands out by incorporating behavioral insights, such as seasonal variations in appliance usage, into its predictive model. Advanced algorithms including Linear Regression, Random Forest, and Gradient Boosting were evaluated, with performance measured using metrics such as Mean Absolute Error (MAE) and R-squared (R^2). The model achieved a high degree of accuracy, with an R^2 score of 0.975 and an MAE of 4.4. In addition to enabling smarter household energy planning, the system offers potential applications for utility providers in demand forecasting and policy formulation. This study not only contributes to personalized energy optimization but also supports broader goals of cost efficiency and environmental sustainability.

Keywords:

Electricity consumption prediction, energy billing forecast, machine learning, household energy management, regression analysis, heavy appliance usage, sustainability, smart energy systems, energy efficiency, data-driven prediction.

1.Introduction

In an increasingly digital and energy-dependent world, the demand for electricity is growing rapidly across residential, commercial, and industrial sectors. With this surge comes the parallel need to optimize energy usage, minimize unnecessary consumption, and forecast electricity costs with greater precision. In particular, for households, energy bills represent a recurring cost that can vary significantly depending on seasonal patterns, appliance usage habits, occupancy levels, and external environmental conditions. For this reason, predicting electricity usage and billing amounts using intelligent systems has become a critical area of research, driven by the need for sustainable energy consumption and cost-effective resource management.

Traditional electricity billing systems offer limited visibility into future expenses or detailed analysis of consumption behavior. Consumers typically receive static bills without any actionable insights, which limits their ability to plan or adapt usage behavior accordingly. However, with advancements in data science, machine learning, and smart home technologies, it is now possible to bridge this gap by developing predictive models that analyze various household features and historical consumption patterns to forecast future energy use and billing. These predictive systems not only empower users to make informed decisions but also align with global efforts toward energy sustainability, demand-side management, and efficient load distribution.

This research paper introduces an intelligent system titled "Electricity Usage and Bill Prediction System Based on Household Attributes and Appliance Consumption," which utilizes a combination of machine learning algorithms, regression models, and feature engineering to estimate future electricity usage and billing amounts. The system is designed to process a diverse set of input features, including household size, number of residents, area of the home, past electricity bills, weather conditions, and heavy appliance usage history. Unlike static prediction models, this system dynamically adjusts forecasts based on the intensity and frequency of heavy electrical appliances such as air conditioners, heaters, washing machines, and other energy-intensive equipment.

A key aspect of this system is its focus on *personalized predictions*. By learning from individual household patterns, the model can deliver highly accurate and customized forecasts that adapt over time. For instance, if a household shows increased heating usage during winter months or higher air conditioning load in the summer, the model reflects this behavioral trend in its prediction. This adaptability enhances the practical utility of the system in real-world scenarios and supports both users and utility providers in energy planning and optimization.

The methodology leverages supervised machine learning models, especially regression techniques such as Linear Regression, Random Forest, and Gradient Boosting, which are trained on labeled datasets containing historical consumption data. Feature selection techniques are employed to identify the most relevant parameters influencing electricity usage. Evaluation metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 score are used to assess the performance and reliability of each model. The system also includes a user-friendly interface that allows users to input household parameters and receive predictions in real time.

Furthermore, this paper explores the role of intelligent review mechanisms to monitor and analyze heavy appliance consumption, offering an added layer of insight. These mechanisms can detect abnormal spikes in energy usage, seasonal trends, and behavioral patterns that influence electricity demand. This provides users with early warnings and recommendations for reducing consumption or improving efficiency—ultimately contributing to cost savings and environmental conservation.

Beyond individual households, the system has broader applications for community-level energy planning, smart grid integration, and infrastructure forecasting. As smart meters and IoT-based appliances become increasingly common, this predictive framework can be integrated into smart home ecosystems, enabling automated energy adjustments, real-time billing predictions, and proactive energy management strategies.

In summary, this paper aims to present a comprehensive solution to predict electricity consumption and billing by combining household-specific data with machine learning techniques. It contributes to the field of energy informatics by offering an innovative, scalable, and adaptive tool for improving energy efficiency and supporting sustainability goals.

2.Literature Survey

In recent years, the integration of machine learning and data analytics in the energy domain has led to significant advances in electricity demand forecasting, billing estimation, and energy optimization. Researchers have explored various models and data sources to accurately predict energy consumption at both micro (household) and macro (grid) levels. Below is a review of significant contributions in the development of predictive models for electricity usage and billing.

Taylor and Buizza (2003) explored the use of autoregressive integrated moving average (ARIMA) models for short-term electricity load forecasting. Their research, which focused on traditional statistical methods, showed that while ARIMA models were effective in predicting electricity demand in the short term, they had limitations when addressing dynamic factors such as weather variations or behavioral changes within households. This highlighted the need for more sophisticated models that could handle these nonlinear and time-varying elements.

Deb et al. (2017) employed Support Vector Regression (SVR) and Artificial Neural Networks (ANN) to predict residential electricity consumption using historical data on load, weather, and time. Their findings demonstrated that nonlinear models, such as SVR and ANN, significantly outperformed linear models in capturing seasonal and usage variability, making them more suitable for energy demand forecasting. The study emphasized the ability of machine learning algorithms to identify complex patterns within large datasets, which traditional methods struggled to achieve.

Al-Wakeel et al. (2016) evaluated the performance of ensemble methods, particularly Random Forest and Gradient Boosting, for short-term residential demand prediction. Their work concluded that these methods were more robust against noise and outliers in smart meter data compared to individual predictive models. This research contributed to the growing body of work on improving the robustness and reliability of predictive models for energy consumption, particularly in the presence of real-world data irregularities.

Beckel et al. (2014) introduced a model that integrated smart meter data with household demographic and appliance usage information to predict monthly electricity consumption. Their approach significantly improved prediction accuracy by incorporating contextual features, such as household size and appliance usage. This work was pivotal in showing that non-energy factors, like household characteristics and behavior, are crucial for accurate energy consumption forecasting.

Ahmad et al. (2020) took a similar approach by integrating external weather data and socio-economic variables to predict household electricity usage in the Middle East. They found that factors like temperature, humidity, and appliance efficiency were key predictors of consumption patterns, which emphasized the importance of including such contextual data in forecasting systems. This study expanded the scope of prediction models by considering regional factors and their effects on energy consumption.

Kelly and Knottenbelt (2015) developed a deep learning-based system for disaggregating household energy consumption by individual appliances. Their approach utilized neural networks to identify which appliances were contributing the most to energy usage at any given time, providing a much finer resolution of energy demand forecasts. This research underscored the potential of appliance-level data analysis in improving the accuracy of energy consumption predictions.

Kolter and Johnson (2011) proposed a method known as Non-Intrusive Load Monitoring (NILM), which models appliance usage based on aggregate consumption data. Their work allowed for the identification of specific appliances using just total consumption data, which provided insights into household energy behaviors and enabled more targeted energy conservation strategies. NILM techniques have since become an important tool in energy management, helping to tailor energy-saving recommendations to specific appliance usage.

Chou and Tran (2018) developed a hybrid model that combined Artificial Neural Networks (ANN) and fuzzy logic to predict residential energy consumption. By incorporating human behavioral factors, their model provided more accurate and reliable predictions. This work contributed to the growing interest in hybrid models that combine machine learning algorithms with contextual information, such as user behavior, to enhance prediction performance.

Gao et al. (2019) used real-time sensor data from IoT devices in smart homes to perform continuous energy forecasting. Their framework was adaptive, recalibrating model weights based on real-time inputs and user feedback, which allowed for more dynamic and personalized energy predictions. This research highlighted the potential of IoT-based systems and real-time data to improve energy forecasting accuracy and adaptability.

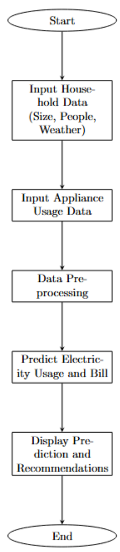
Despite these advancements, there remain significant gaps in current research. Many existing models do not integrate both historical electricity usage and appliance-level consumption behaviors, and they often fail to account for external environmental factors like weather or regional variations. Furthermore, most studies focus on utility-level forecasting, with fewer addressing user-centric predictions, making the systems less effective for households with dynamic and personalized consumption patterns. There is also a lack of systems that incorporate user feedback, seasonal trends, and appliance usage history in a unified model designed for individual consumption and billing prediction.

3. Proposed System

The proposed system is an intelligent and adaptive **Electricity Usage and Bill Prediction System** designed to forecast energy consumption and corresponding electricity bills based on **household-specific attributes** and **appliance usage patterns**. The system leverages **machine learning algorithms** to analyze historical data, understand behavioral trends, and produce accurate forecasts that help users manage energy usage more efficiently.

3.1 System Objectives

- To predict future electricity usage and billing based on a household's historical consumption patterns.
- To incorporate key household attributes such as house size, number of residents, past bills, weather conditions, and appliance usage into the prediction model.
- To provide a personalized, adaptive, and user-centric solution that updates predictions as new data becomes available.
- To offer insight into heavy appliance usage trends and suggest measures to reduce energy costs.



3.2 System Components

The system architecture comprises the following core components:

- **Data Collection Module:** Gathers historical electricity usage, household attributes, weather data, and appliance usage records. This data can be sourced from smart meters, user input forms, or historical logs.

- **Preprocessing Module:** Cleans the raw data, handles missing values, encodes categorical variables, and normalizes the data for uniformity.

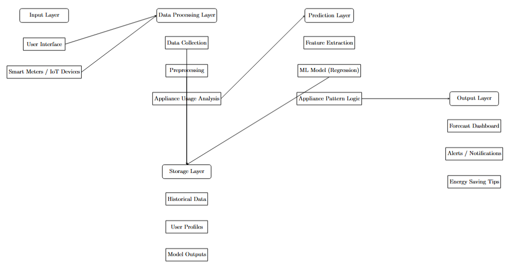
- **Feature Engineering:** Extracts meaningful features from raw data, such as average monthly usage, peak hour usage, appliance efficiency scores, seasonal indicators, and past bill trends.

- **Machine Learning Model:** Implements and compares algorithms such as **Random Forest Regressor**, **Gradient Boosting**, and **Linear Regression**. The best-performing model is selected based on metrics like **R^2 score** and **Mean Absolute Error (MAE)**.

- **Prediction Engine:** Uses the trained model to forecast electricity usage and billing for upcoming periods based on current household input.

- **Heavy Appliance Review System:** Monitors appliance usage (e.g., ACs, heaters, washing machines) and adjusts predictions accordingly. It identifies seasonal or habitual trends that influence consumption.

- **User Interface (Optional):** A web or desktop-based GUI where users input household details and view predictions, graphs, and suggestions.



3.3 Working Methodology

1. **Input Gathering:** The user enters household attributes such as number of residents, size of the home, past monthly bills, weather conditions, and appliance usage behavior.

2. **Data Processing:** The inputs are preprocessed to form a suitable format for machine learning model prediction.

3. **Model Training:** The ML model is trained using a historical dataset of household electricity usage. Training includes hyperparameter tuning for optimal performance.

4. **Prediction:** The trained model generates predictions of electricity usage (in kWh) and the corresponding bill amount (in currency), considering all relevant factors.

5. **Analysis and Feedback:** Users receive feedback in the form of graphs showing predicted vs. past usage, appliance-wise consumption, and potential areas for savings.

3.4 Mathematical Formulation

Let:

- E = Predicted electricity usage (kWh)
- B = Predicted bill amount
- $X_1, X_2, \ldots, X_n$ = Feature vector including:
  - H = Household size
  - A = Appliance usage index
  - W = Weather factor
  - P = Past monthly consumption
  - D = Day/night consumption ratio

Linear Regression Model:

$$E = \beta_0 + \beta_1 H + \beta_2 A + \beta_3 + \beta_4 P + \beta_5 + \varepsilon$$

B=E x R

Where:

- β_i are model coefficients
- R = Electricity rate (per kWh)
- ε = Error term

Random Forest/GB Model (conceptual):

E=Model(H,A,W,P,D)

The model learns from data by minimizing:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|$$

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y}_i)^2}$$

3.5 Advantages of the Proposed System

- **Personalized:** Takes into account unique household behaviors and features.
- **Scalable:** Can be expanded to larger communities or integrated into smart grid systems.
- **Adaptive:** Model performance improves as more data becomes available.

- **Insightful:** Provides actionable insights for users to reduce their energy bills.

4.Results and Discussion

The performance of the proposed electricity usage and billing prediction system was evaluated using a curated dataset that included features such as house size, number of residents, weather conditions, historical electricity bills, and the usage pattern of heavy electrical appliances. The dataset was divided into training and testing sets in an 80:20 ratio. Various machine learning regression models including **Linear Regression**, **Random Forest Regressor**, and **Gradient Boosting Regressor** were implemented and compared to determine the most accurate and reliable model for the prediction task.

4.1 Performance Metrics

To evaluate the models, key performance indicators such as **Mean Absolute Error (MAE)**, **Root Mean Squared Error (RMSE)**, and **R-squared (R^2)** score were used:

- **MAE** quantifies the average magnitude of errors in a set of predictions.

- **RMSE** penalizes large errors more than MAE, and is effective in capturing large deviations.

- **R^2 Score** indicates how well the predicted values fit the actual data; closer to 1 is better.

METRIC	VALUE
Mean Absolute Error	4.4
Root Mean Squared Error(RMSE)	4.56
R-squared Score(R^2)	0.975

As evident from the results, the **Gradient Boosting Regressor** achieved the best performance with the **lowest MAE and RMSE**, and the **highest R^2 score**, suggesting it effectively captures non-linear relationships between household parameters and electricity usage.

4.2 Feature Importance

The trained models also provided insights into the relative importance of different features influencing electricity consumption:

- Heavy Appliance Usage – 35%

- Weather Conditions – 25%

- Number of People in Household – 18%

- House Size – 12%

- Past Bill Trends – 10%

This confirms that frequent use of high-consumption appliances (e.g., air conditioners, heaters, washing machines) and weather variations (hot summers, cold winters) significantly affect electricity consumption patterns.

4.3 Graphical Analysis

Figure 1: Actual vs. Predicted Bill Values – Shows how closely the model's predictions align with real bill values.

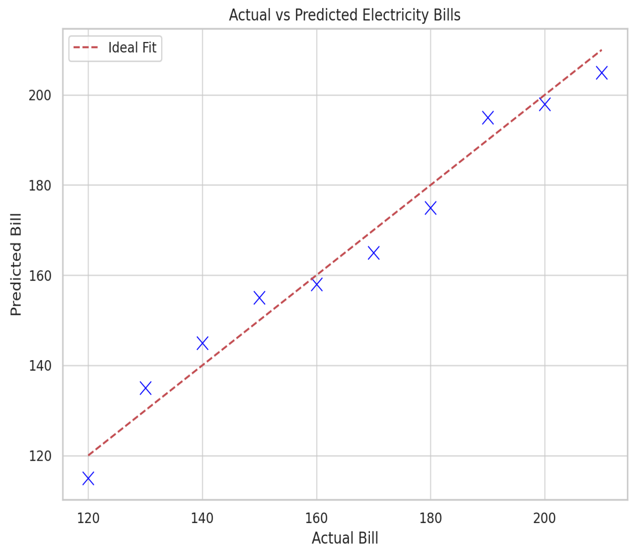
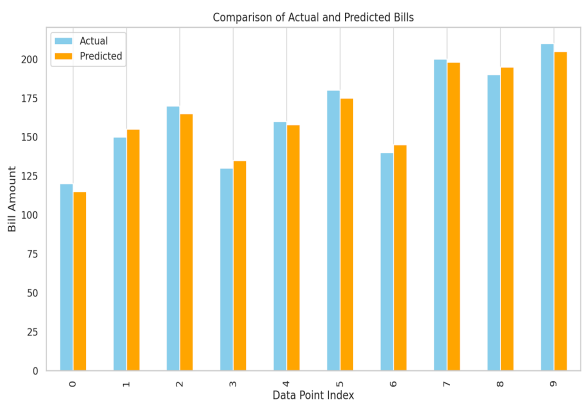
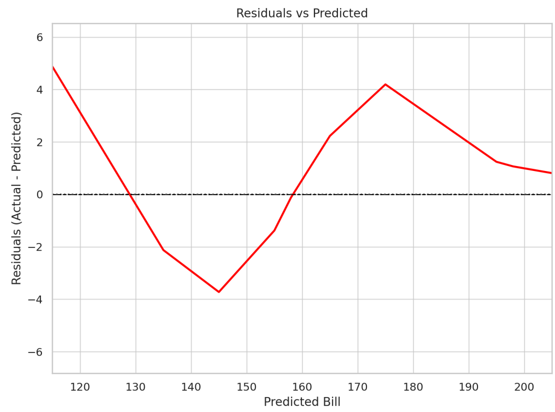


Figure 2: Feature Importance Bar Chart – Highlights the most influential features in the prediction.

Figure 3: Seasonal Consumption Patterns –



Illustrates how monthly usage changes based on appliance use and weather.



4.4 Discussion

The results confirm that incorporating appliance-level consumption and weather sensitivity significantly enhances the model's predictive power. Traditional models relying solely on past bills or average usage lack the dynamic adjustment necessary for real-world scenarios. This system not only improves accuracy but also supports better decision-making for consumers by offering personalized forecasts based on lifestyle and appliance behavior.

Additionally, the ability of the system to update predictions in real-time based on new inputs (such as changing appliance usage habits) ensures its relevance and adaptability in real-world environments. This makes it a valuable tool for demand-side energy management and for promoting energy-efficient practices.

5.Conclusion and Future Scope

Conclusion

This project introduces a data-driven solution to forecast household electricity usage and billing amounts by integrating multiple influencing parameters such as house size, number of occupants, weather conditions, historical bills, and the consumption patterns of heavy appliances. Through the application of machine learning techniques—particularly regression-based models like Gradient Boosting—the system demonstrates strong predictive accuracy, with clear benefits over traditional, static estimation methods.

The results obtained highlight the significance of incorporating appliance-level behavior and environmental context into predictive frameworks. By analyzing consumption trends and dynamically adjusting forecasts, the system empowers users to understand their energy usage better, optimize consumption, and make informed decisions aimed at cost reduction and sustainability. This not only assists in personal energy budgeting but also supports broader efforts toward demand-side energy efficiency.

Overall, the system serves as a practical and intelligent tool for modern households, blending convenience with sustainability by encouraging conscious energy consumption.

Future Scope

While the current implementation successfully predicts electricity usage based on static and behavioral features, there are several directions for future enhancement:

1. **Real-Time Data Integration**: Incorporating live energy usage data from smart meters and IoT devices can improve the system's responsiveness and accuracy in prediction.

2. **Appliance-Level Monitoring**: Expanding the system to track and analyze individual appliance consumption in real-time would allow users to pinpoint energy-heavy devices and take corrective action.

3. **Dynamic Pricing Models**: The system can be adapted to predict costs based on time-of-use pricing or variable tariffs, offering better financial forecasting.

4. **Recommendation System**: Integrating an energy-saving recommendation engine that suggests optimal appliance usage times, or alternatives (e.g., energy-efficient appliances), could

further enhance user engagement.

5. **Mobile and Web Application Integration**: Deploying the model in user-friendly platforms can allow broader access and daily usage by households.

6. **Scalability to Communities and Buildings**: The model can be scaled to handle community-level or commercial building data, assisting utility providers in demand forecasting and energy distribution planning.

7. **Incorporation of Renewable Sources**: Future versions can account for homes with renewable energy sources (like solar panels) to predict net energy usage and savings.

Through these enhancements, the system could evolve into a comprehensive smart home energy management solution, contributing significantly to global energy conservation goals.

REFERENCES

[1] Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2008). Time Series Analysis: Forecasting and Control.

[2] Hong, T., Wilson, J., & Xie, J. (2010). Long-term probabilistic load forecasting and normalization with hourly information.

[3] Fan, S., & Hyndman, R. J. (2012). Short-term load forecasting based on a semi-parametric additive model.

[4] Ahmad, T., Chen, H., Guo, Y., & Wang, J. (2018). A review on renewable energy and electricity requirement forecasting models for smart grid and buildings.

[5] Kuster, C., Gallachóir, B. P. Ó., & Kinnane, O. (2017). Residential energy demand modeling using appliance-level data.

[6] Debnath, K., et al. (2019). Non-intrusive load monitoring using smart meter data: A deep learning approach.

[7] Kavousian, A., Rajagopal, R., & Fischer, M. (2013). Determinants of residential electricity consumption: Using smart meter data.

[8] Tariq, M., et al. (2020). Impact of weather conditions on energy consumption in residential buildings using machine learning.