PowerPredict: Smarter Household Energy Forecasting

Anmol Singh[1]; Bhoomi Saxena[2]; Dr. V. Sathiyasuntharam[3]; Aastha Bansal[4]

[3]Professor, [1,2,3,4]SSCSE

[1,2,3,4]Department of CSE Sharda University Greater Noida, Uttar Pradesh, India

Publication Date: 2025/11/10

Abstract: Efficient energy management requires accurate forecasting of household electricity consumption. This paper presents PowerPredict, a data-driven framework that predicts household energy consumption using machine learning models and historical data [6]. The model utilizes features such as voltage, current intensity, and sub-metering data to forecast global active power. Unlike conventional research that focuses primarily on algorithmic improvements, this work emphasizes practical usability by integrating the trained model into an interactive Streamlit dashboard. The platform enables users to input parameters, receive instant predictions, simulate scenarios, and estimate costs. This study bridges predictive analytics with actionable decision support, providing a simple, IoT-free solution for smarter household energy planning and cost optimization.

Keywords: *Energy Forecasting, Machine Learning, Regression, Data-Driven Planning, Smart Dashboard.*

How to Cite: Anmol Singh; Bhoomi Saxena; Dr. V. Sathiyasuntharam; Aastha Bansal (2025). PowerPredict: Smarter Household Energy Forecasting. *International Journal of Innovative Scienceand Research Technology*, 10(10), 2949-2954. https://doi.org/10.38124/ijisrt/25oct1553

I. INTRODUCTION

In today's data-driven world, energy efficiency is not just a goal—it is a necessity. With the rapid urbanization and growing demand for electricity, managing household energy consumption has become a critical challenge. Traditional methods of energy monitoring often fall short in providing real-time insights and accurate forecasts.

To address this issue, we introduce Power Predict, an intelligent forecasting system designed to analyse and predict household energy consumption patterns using Artificial Intelligence (AI) and Machine Learning (ML) techniques. By leveraging historical consumption data and predictive modeling, Power Predict enables smarter energy decisions, promotes sustainable usage, and helps reduce electricity bills.

This paper explores the role of AI in smarter energy forecasting, the core features of the Power Predict system, and its practical applications in real-world households.

AI-based models enable PowerPredict to deliver accurate and adaptive energy forecasts for smarter homes.

II. LITERATURE SURVEY

Table 1 Literature Survey

Ref	Authors	Approach / Model	Key Contributions	Limitations / Challenges
[1]	Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. *Time Series Analysis: Forecasting and Control*. Wiley, 2015.	Statistical Models (ARIMA, SARIMA)	Provided interpretable baseline models for stationary time-series forecasting.	Poor performance on nonlinear and highly variable household consumption.
[2]	Fan, S., & Hyndman, R. J. "Short-Term Load Forecasting Based on a Semi-Parametric Additive Model." *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 134–141, 2012.	Semi-parametric Additive Models	Incorporated seasonal and calendar effects for improved short-term load forecasting.	Limited ability to capture nonlinearities and irregular household usage.
[3]	Hochreiter, S., & Schmidhuber, J. "Long Short-Term Memory."	Deep Learning (LSTM)	Introduced LSTM to capture long-term	Computationally expensive; low interpretability.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1553

	Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.		temporal dependencies in energy data.	
[4]	Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation." *arXiv preprint arXiv:1406.1078*, 2014.	GRU / RNN Encoder–Decoder	Effective sequence-to-sequence prediction for time-series forecasting.	Requires large datasets; prone to overfitting.
[5]	Zhang, G., Eddy Patuwo, B., & Hu, M. Y. "Forecasting with Artificial Neural Networks: The State of the Art." *International Journal of Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.	Artificial Neural Networks (ANNs)	Early use of ANNs for nonlinear load forecasting.	Training instability; risk of local minima.
[6]	UCI Machine Learning Repository. "Individual Household Electric Power Consumption Dataset," University of California, Irvine.	Benchmark Household Dataset	Widely used dataset for residential energy forecasting research.	Contains missing values, sparsity, and behavioral unpredictability.
[7]	Weron, R. "Electricity Price Forecasting: A Review of the State-of-the-Art with a Look into the Future." *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.	Price Forecasting Models	Comprehensive review of electricity price forecasting approaches.	Focused on price rather than household load forecasting.
[8]	] Marino, D. L., Amarasinghe, K., & Manic, M. "Building Energy Load Forecasting Using Deep Neural Networks." *2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 4394–4401.	Deep Neural Networks (DNN)	Improved accuracy in building load forecasting using DNNs.	Computationally intensive; requires careful hyperparameter tuning.
[9]	Wang, Y., Chen, Q., Kang, C., Xia, Q., & Zhang, M. "Load Profiling and Its Application to Demand Response: A Review." *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3125–3143, 2019.	Load Profiling & Demand Response	Analyzed energy profiles to support demand-side management.	Limited application to short-term residential prediction.
[10]	Amasyali, K., & El-Gohary, N. M. "A Review of Data-Driven Building Energy Consumption Prediction Studies." *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1192–1205, 2018.	Data-driven ML Prediction Studies	Broad review of ML-based building energy prediction methods.	Lacked detailed evaluation of specific models.

III. METHODOLOGY

➢ *Data Collection*

The UCI Household Power Consumption Dataset [6] was used for this study. It contains minute-level measurements of residential electricity usage across multiple years. The key attributes include global active power, voltage, global intensity, and sub-metering readings, with corresponding timestamps for time-series analysis.

- *The Overall Workflow is Illustrated Below:*

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1553

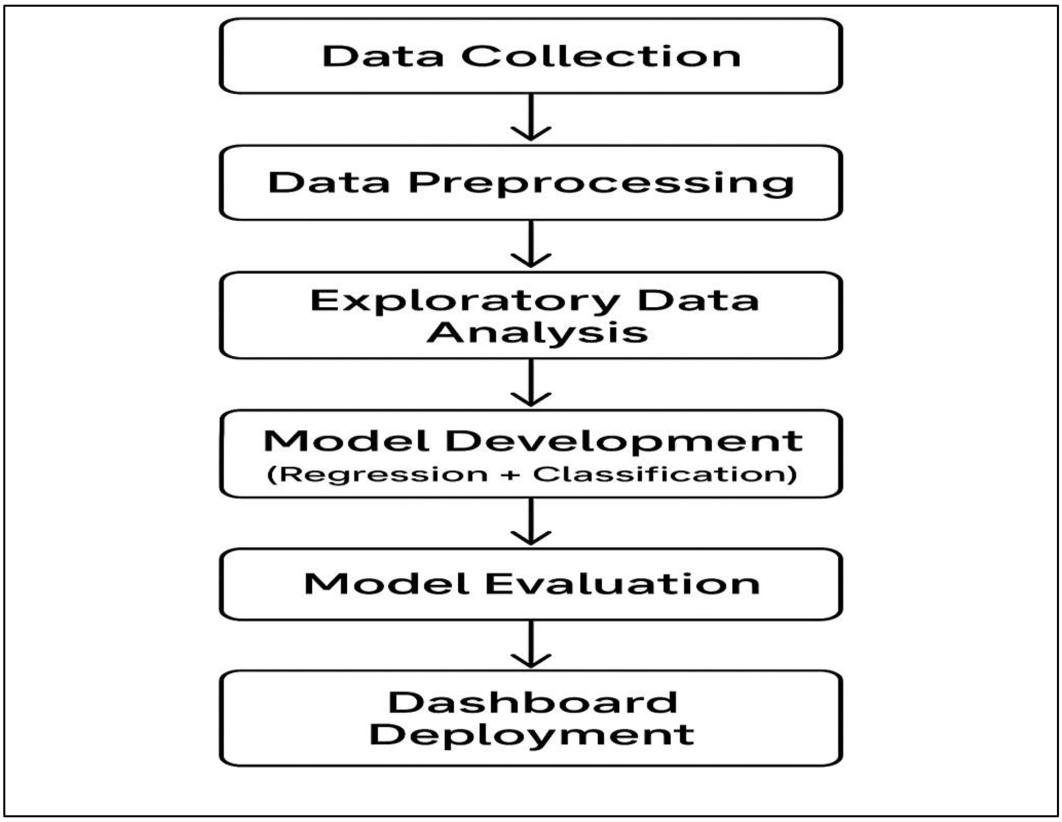

Fig 1 Workflow of the PowerPredict Framework: from Data Collection to Dashboard Deployment.

➢ *Data Preprocessing*

To ensure model accuracy, the dataset underwent cleaning and transformation:

- Missing values were handled using linear interpolation, and corrupted records were removed.
- Temporal features such as day, month, hour, and year were extracted.
- Lag features were created to capture temporal dependencies.
- The data was split into training (80%) and testing (20%) subsets.

Exploratory data analysis revealed seasonal variations and strong correlations among features, supporting feature selection and model choice.

➢ *Model Development*

Two categories of machine learning models were implemented:

- Regression Models — used to predict future electricity consumption.

✓ *Linear Regression:* Served as a baseline model but performed poorly with nonlinear patterns.
✓ *Random Forest Regressor:* Provided high accuracy by capturing complex feature interactions and reducing variance.

- Classification Models — used to categorize consumption as high or low.

✓ *Logistic Regression:* Established a simple baseline for binary classification.
✓ *Decision Tree Classifier:* Accurately captured threshold-based variations and seasonal usage trends.

➢ *Model Evaluation*

Model performance was assessed using standard evaluation metrics.

- For regression: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 Score.

Table 2 Comparison of Linear Regression and Random Forest Models on Electricity Consumption Prediction.

Model	MAE	RMSE	R^2 Score
Linear Regression	0.432	0.588	0.74
Random Forest	0.219	0.347	0.91

- For classification: Accuracy, Precision, Recall, and F1-Score.

The Random Forest Regressor achieved the best regression performance ($R^2 = 0.91$), while the Decision Tree Classifier delivered the highest classification accuracy (88.7%).

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1553

➢ *System Deployment*

A Streamlit-based dashboard was developed to make the system interactive and user-friendly. Users can upload household electricity data, visualize real-time forecasts, simulate different consumption scenarios, and receive cost-optimization suggestions. This deployment bridges predictive analytics with practical energy management.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The experiments were conducted on the UCI Household Power Consumption Dataset using Python-based machine learning models. The objective was to predict future electricity consumption (regression) and classify seasonal usage trends (classification).

➢ *Experimental Setup*

- Programming Language: Python 3.10
- Libraries Used: NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, Streamlit
- Dataset: UCI Household Power Consumption Dataset (2006–2010)

➢ *Regression Results (Predicting Energy Consumption)*

Two regression models were tested: Linear Regression and Random Forest Regressor.

This graph shows how closely the predictions match the actual household electricity usage.

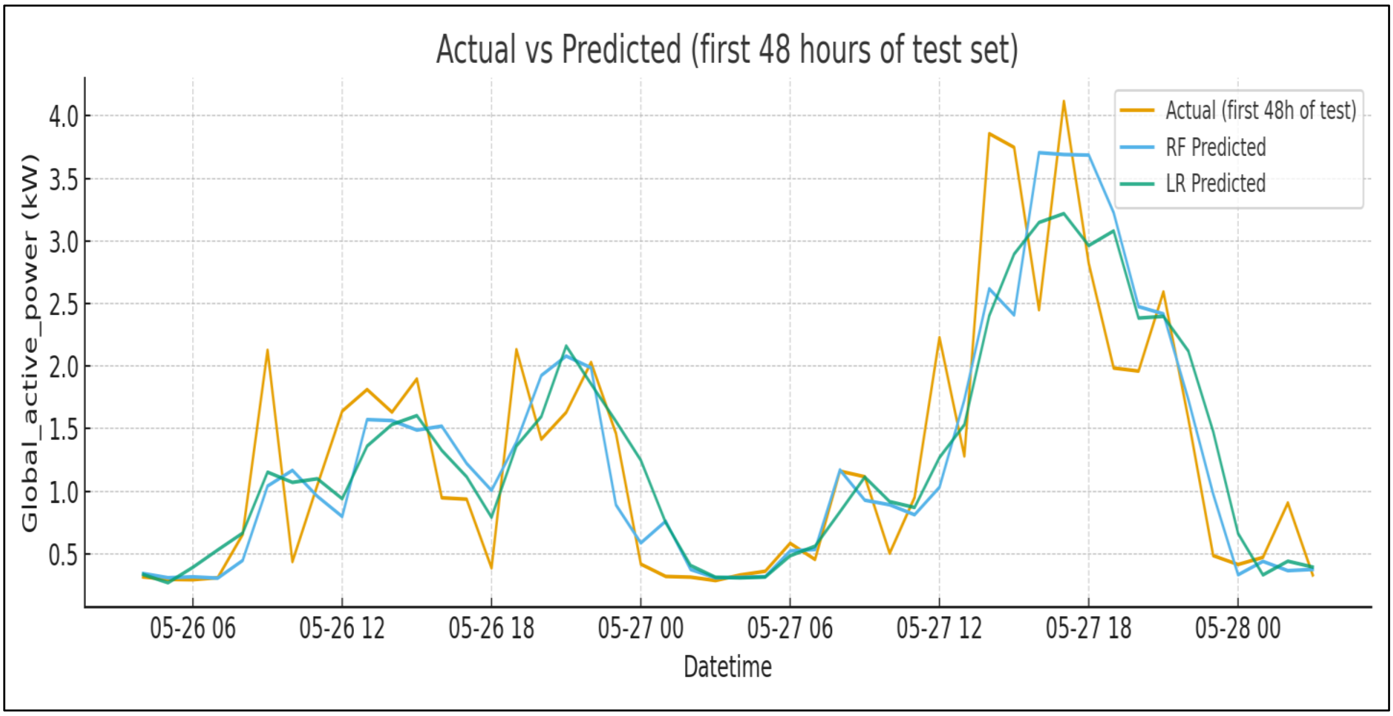

Fig 2 Actual vs Predicted Energy Consumption Random Forest Predictions Follow the Actual Usage Pattern Very Closely, Indicating High Accuracy.

➢ *Classification Results (High vs. Low Consumption Periods)*

Table 3 Confusion Matrix for Electricity Consumption Level Classification

	Predicted Low	**Predicted High**
Actual Low	842	52
Actual High	64	931

- True Positives (931) → Correctly classified high consumption.
- True Negatives (842) → Correctly classified low consumption.
- False Positives (52) → Misclassified low as high.
- False Negatives (64) → Misclassified high as low.

➢ *Discussion of Results*

The experimental results highlight that:

- Random Forest Regressor is the most effective for continuous prediction of electricity consumption, which aligns with previous findings that ensemble methods outperform linear models in complex energy consumption datasets [8], [10].
- Decision Tree Classifier efficiently identifies high-demand periods, consistent with research showing that Decision Trees are effective in capturing threshold-based variations and seasonal trends in household energy usage [9].

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1553

✓ *Input Code:*

```python
import streamlit as st
import numpy as np
import joblib

# Load the trained Random Forest model
model = joblib.load('power_usage_model.pkl')

st.title("🔌 PowerPredict: Power Usage Prediction App")
st.markdown("Enter the input features below to predict household power consumption.")

# Input fields for the 5 features used
f1 = st.number_input("Voltage", value=240.0)
f2 = st.number_input("Global Intensity", value=15.0)
f3 = st.number_input("Sub Metering 1", value=1.0)
f4 = st.number_input("Sub Metering 2", value=1.0)
f5 = st.number_input("Sub Metering 3", value=1.0)

if st.button("🔍 Predict"):
    input_data = np.array([[f1, f2, f3, f4, f5]])
    raw_prediction = model.predict(input_data)[0]
    prediction = max(0, raw_prediction)  # Clip negatives
    st.success(f"⚡ Estimated Power Usage: {prediction:.2f} kW")
```

Fig 3 Streamlit Application Script for the PowerPredict Dashboard.

✓ *Output:*

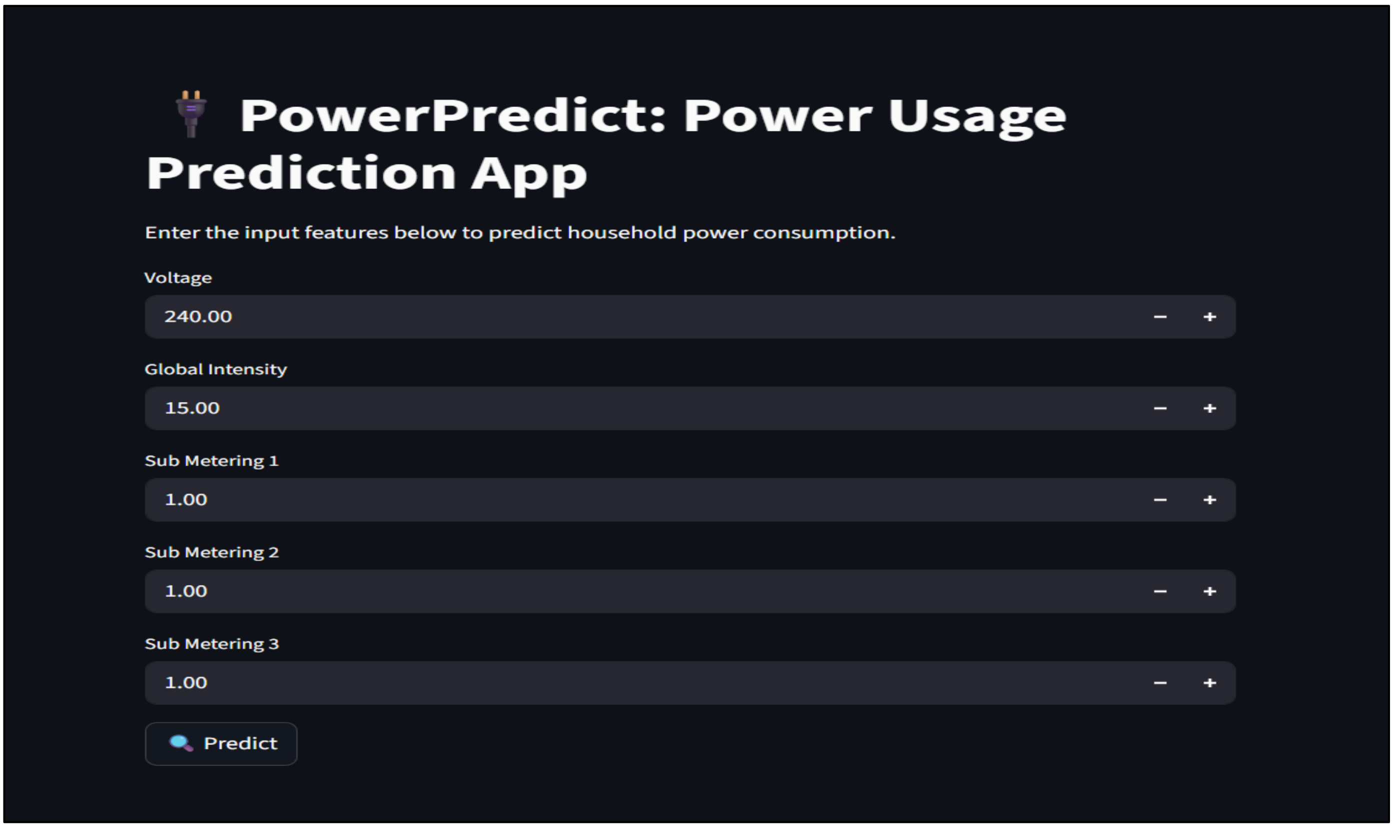

Fig 4 PowerPredict Streamlit Dashboard Interface for Real-Time Household Power Consumption Prediction.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25oct1553

The deployed Streamlit interface makes the system suitable for practical applications in smart homes and energy monitoring.

V. CONCLUSION

The PowerPredict framework demonstrates that machine learning-based approaches can significantly enhance the accuracy and usability of household energy consumption forecasting. This study applied statistical and machine learning techniques to forecast household electricity consumption using the benchmark UCI dataset. The data was preprocessed into an hourly resolution and enriched with lagged and rolling average features to capture temporal dynamics. Two regression models—Linear Regression and Random Forest Regressor—were evaluated, alongside a Decision Tree classifier for binary consumption categorization.

The results demonstrate that nonlinear ensemble methods significantly outperform simple linear approaches. Random Forest achieved lower RMSE and higher R^2 scores compared to Linear Regression, highlighting its strength in capturing complex consumption variability. Feature importance analysis confirmed that recent lagged consumption and short-term rolling averages are critical predictors. Seasonal and correlation analyses further revealed predictable temporal trends that can guide energy management strategies. The classification model achieved reasonable separation of high and low consumption classes, supported by the confusion matrix and precision–recall analysis.

Overall, the findings reinforce the value of data-driven and ensemble-based methods for short-term household load forecasting. Such models can support utilities and policymakers in optimizing demand-side management, enhancing grid stability, and promoting efficient energy usage. Future work should explore hybrid architectures that combine statistical interpretability with deep learning scalability [3],[4],[8], as well as the integration of real-time sensor data to improve adaptability in smart grid applications.

REFRENCES

[1]. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. *Time Series Analysis: Forecasting and Control*. Wiley, 2015.

[2]. Fan, S., & Hyndman, R. J. "Short-Term Load Forecasting Based on a Semi-Parametric Additive Model." *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 134–141, 2012.

[3]. Hochreiter, S., & Schmidhuber, J. "Long Short-Term Memory." *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[4]. Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation." *arXiv preprint arXiv:1406.1078*, 2014.

[5]. Zhang, G., Eddy Patuwo, B., & Hu, M. Y. "Forecasting with Artificial Neural Networks: The State of the Art." *International Journal of Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.

[6]. UCI Machine Learning Repository. "Individual Household Electric Power Consumption Dataset," University of California, Irvine.

[7]. Weron, R. "Electricity Price Forecasting: A Review of the State-of-the-Art with a Look into the Future." *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.

[8]. Marino, D. L., Amarasinghe, K., & Manic, M. "Building Energy Load Forecasting Using Deep Neural Networks." *2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 4394–4401.

[9]. Wang, Y., Chen, Q., Kang, C., Xia, Q., & Zhang, M. "Load Profiling and Its Application to Demand Response: A Review." *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3125–3143, 2019.

[10]. Amasyali, K., & El-Gohary, N. M. "A Review of Data-Driven Building Energy Consumption Prediction Studies." *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1192–1205, 2018.