

Exploring The Relationship Between Weather Patterns and Energy Consumption in Smart Homes: A Regression Analysis

Pravalika[1], Bobbala Navya[2], Gandham Bhavani[3], Azmeera Nagaraju[4]

[1]Department of Computer Science and Engineering, Kommuri Pratap Reddy Institute of Technology, Ghatkesar, Hyderabad, Telangana.

[2] UG Scholar, Department of Computer Science and Engineering, Kommuri Pratap Reddy Institute of Technology, Ghatkesar, Hyderabad, Telangana.

ABSTRACT

The history of studying weather's impact on energy consumption dates back to the early days of modern energy systems. Historically, energy demand was primarily analyzed based on seasonal variations and historical consumption data. With the digital revolution, the integration of weather data into energy analysis began to gain prominence. Early studies used basic statistical models to correlate weather patterns with energy usage. However, the emergence of machine learning techniques in the last two decades has revolutionized this field. The utilization of decision trees, random forests, and neural networks has enabled researchers to create highly accurate predictive models. This project builds upon this historical evolution, leveraging cutting-edge technologies to delve deeper into the relationship between weather patterns and energy consumption in the context of smart homes, contributing to the ongoing evolution of energy-efficient technologies and practices. Thus, this research aims to investigate the intricate relationship between weather patterns and energy consumption in smart homes through a regression analysis. Leveraging machine learning techniques, the study explores predictive models to comprehend how weather variables impact the total energy load in these environments. The analysis involves the use of decision tree and random forest regression algorithms, providing valuable insights into energy consumption patterns under varying weather conditions.

Keywords: Weather impact, Energy Consumption, Smart Homes, Machine Learning, Predictive models, Regression analysis.

1. INTRODUCTION

The intersection of technology and sustainability has paved the way for innovations in smart home systems, revolutionizing the way we interact with our living spaces. In this era of smart homes, understanding energy consumption patterns is pivotal not only for homeowners seeking efficient energy management but also for energy providers and policymakers striving for sustainable practices. The concept of smart homes has evolved significantly with the advent of the Internet of Things (IoT) and artificial intelligence. Today, smart homes are equipped with an array of sensors and devices that collect vast amounts of data, including temperature, humidity, occupancy, and energy usage. Smart meters are used to accurately record the amount of electricity consumption at a very high frequency, dramatically changing the collection of electricity data and driving the household energy transition [1]. High frequency interval meter data, typically hourly and 15 min, provides important and rich information about household consumption patterns. Smart meter data can be used to cluster, classify, predict, and optimize electricity consumption patterns through a series of analytical methods and techniques [2]. The popularity of smart meters has grown rapidly over the past decade, from <2.5 million smart meters deployed globally in 2007 to ~729.1 million in 2019, an increase of 294 times, with the United States and China accounting for the highest percentage, 85.4% [3]. Smart meters provide utilities with detailed information and enable effective demand side management. Two-way AMI meters, which allow communication capability between electric utilities and customers, have been more prevalent after 2013 [4]. By providing real-time or near real-time electricity data, it supports smart consumption applications based on customer preferences and demand.

This data presents an unprecedented opportunity to analyze and understand the factors influencing energy consumption, particularly the impact of weather patterns. Weather variables such as temperature and humidity have long been known to affect energy demand, but the complexity of these interactions demands sophisticated data analysis methods. Traditional methods often fall short in capturing the nuanced relationships, necessitating the use of machine learning algorithms for precise predictions. This project delves deep into the intricate relationship between weather patterns and energy usage in smart homes. By employing advanced regression analysis and machine learning techniques, the project aims to uncover the underlying patterns, providing valuable insights that can optimize energy consumption, reduce costs, and contribute to a greener future.Understanding the relationship between weather patterns and energy consumption is crucial in the context of smart homes and sustainable energy usage. With the rising importance of smart home technology and the increasing focus on energy efficiency, it is imperative to analyze the factors that influence energy consumption. By deciphering the impact of weather variables such as temperature, humidity, and precipitation on energy load, this study addresses a fundamental need in optimizing energy usage in smart homes. These findings are invaluable for both homeowners and energy providers, enabling them to make informed decisions to enhance energy efficiency and reduce costs.

2. LITERATURE SURVEY

High-frequency electricity data helps understand the electricity consumption patterns in different consumer groups at various time periods, and the changes in behaviors after the adoption of new technologies and demand-side management measures. Further, high-frequency data increases the accuracy of energy consumption forecasts due to the larger variation provided by the data. Applying high frequency electricity data during pandemic times, studies have analyzed and examined the overall impact of COVID-19 on energy consumption and transition in pre- and post-pandemic. The world has seen a shift in people's habits and daily activities due to the pandemic. Therefore, electricity consumption patterns in both residential and commercial buildings have changed. Ku et al. [5] used individual hourly power consumption data within a machine learning framework to examine changes in electricity use patterns due to COVID-19 mandates in Arizona. Chinthavali et al. [6] examined changes in energy use patterns on weekdays and weekends before and after the

Research Paper , **Journal Volume 13, Iss 4, 2023**

COVID-19 pandemic. Raman and Peng [7] used residential electricity consumption data to reveal a strong positive correlation between pandemic progress and residential electricity consumption in Singapore. Li et al. analyzed data from apartments in New York to examine the impact of the number of COVID-19 cases and the outdoor temperature on residential electricity usage [8].

Lou et al. found that the COVID-19 measures increased residential electricity consumption by 4–5% and exacerbated energy insecurity using individual smart meter data from Arizona and Illinois [9]. Sánchez-López et al. explored the evolution of energy demands with hourly data among residential, commercial, and industrial demand during the first wave of COVID-19 [10]. Understanding how household hourly electricity demand changes after the pandemic, especially due to working from home, provides electricity system operators with valuable information in operation and management. Also, based on the changes in the spatial and temporal distributions of energy consumption, policymakers could make better decisions to increase the ratio of power supply from renewable energy sources.

The application of high frequency electricity data could help understand the electricity consumption patterns of specific consumer groups, especially families that have adopted new technologies [e.g., Photovoltaics (PV), batteries, and electric Vehicles (EV)]. Qiu et al. [11] applied a difference-in-differences approach to 1600 EV households' high frequency smart meter data and found that people increased EV charging in lower-priced off-peak hours.

Al Khafaf et al. [12] compared the electricity consumption of consumers with PV and energy storage systems (ESS) against consumers without ESS using over 5,000 energy consumers' 30-min window smart meters recording. They found that on extremely hot days, installing batteries, to some extent, reduces peak power usage in the afternoon. Using household hourly electricity data in Arizona, in [13] Qiu et al. (2022b) found a high degree of heterogeneity in consumption patterns of PV consumers after adding battery storage. As to heat pump adoption, Liang et al. (2022a) provided empirical evidence from Arizona which suggested that heat pumps do not necessarily save energy [14].

3. PROPOSED METHODOLOGY

This research explores the intricate relationship between weather patterns and energy consumption in smart homes, employing sophisticated data analysis techniques and machine learning algorithms. In this endeavour, this work analyzes a dataset containing information about weather variables such as temperature, humidity, and precipitation, alongside energy consumption data from smart homes.

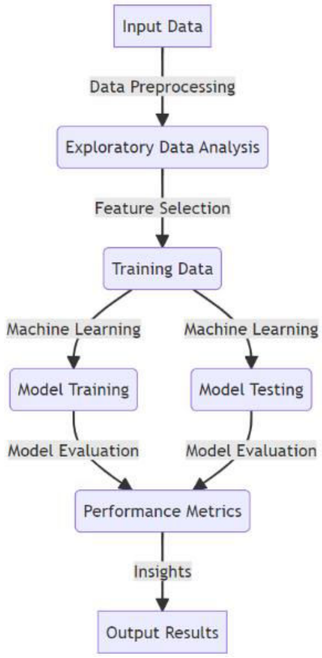

Figure 1: Proposed methodology of ML-based energy consumption prediction in smart homes.

The primary objective is to discern patterns and correlations within this data to understand how weather conditions impact energy usage.

—Data Analysis and Preprocessing: This initiates with data preprocessing, addressing missing values and ensuring data integrity. Basic statistical analyses and visualization tools are employed to gain a comprehensive understanding of the dataset. Exploratory data analysis techniques are utilized to visualize trends, histograms, and correlations among variables, providing valuable insights into the data's structure.

—Machine Learning Models: To uncover the intricate relationships hidden within the data, advanced machine learning models are implemented. The project employs two primary regression algorithms: Decision Tree Regressor and Random Forest Regressor. These algorithms are trained on the preprocessed data, utilizing historical weather and energy consumption patterns to make predictions. Decision trees offer interpretable insights into feature importance, while random forests leverage multiple decision trees for enhanced accuracy and robustness.

—Analysis and Interpretation: The models' predictions are rigorously analyzed, evaluating their accuracy and effectiveness in forecasting energy consumption based on weather patterns. Key performance metrics, such as R-squared scores, are calculated to quantify the models' predictive power. These metrics offer crucial insights into the models' ability to capture the complexities of energy usage dynamics in response to changing weather conditions.

—Significance and Implications: The findings have profound implications for various stakeholders. Homeowners can optimize their energy usage, reducing costs and environmental impact. Energy providers can enhance their demand forecasting, ensuring a stable energy supply. Policymakers gain valuable insights for crafting sustainable energy policies, aligning urban planning with energy efficiency goals. Moreover, the project showcases the potential of machine learning in addressing real-world challenges, underlining its significance in the realm of energy management and sustainability.

—Future Directions: Looking forward, this work lays the foundation for future research avenues. Refining machine learning models, integrating real-time data, exploring regional variations, and diversifying applications across sectors are promising directions. These advancements hold the potential to create even more accurate, responsive, and adaptable energy management systems, ushering in a future of sustainable and efficient energy usage.

Random Forest Regressor

Random Forest is an ensemble learning method, meaning it combines the predictions of multiple individual algorithms (in this case, decision trees) to create a more accurate and robust model. In the case of regression tasks, where the goal is to predict a continuous numerical value (like energy consumption), the algorithm is called a Random Forest Regressor. Overall, the Random Forest Regressor plays a crucial role in accurately predicting energy consumption in smart homes based on weather patterns. Its ability to handle non-linearity, provide feature importance insights, and maintain robustness makes it a suitable choice for this complex predictive task.

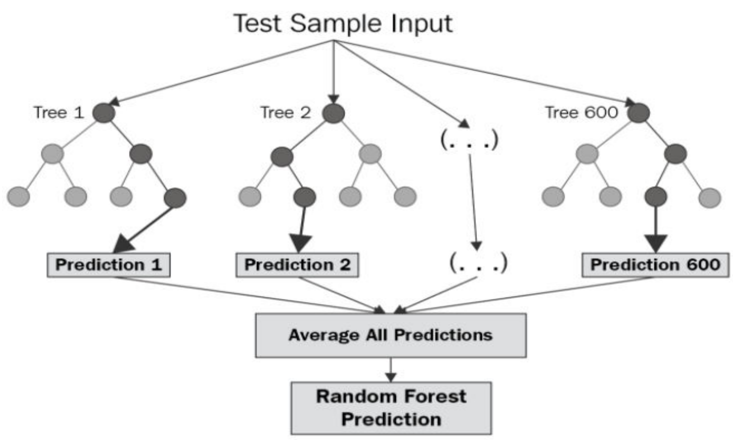

Figure 2: Working of RFR model.

Working

Decision Trees: Random Forest starts by creating a multitude of decision trees. Each tree is trained on a random subset of the data and a random subset of features. This randomness helps the trees to be diverse and not overly reliant on a specific subset of the data.

Voting Mechanism: When it's time to make a prediction, each individual tree in the forest produces its own prediction. In the case of regression, the predictions from each tree are averaged to produce the final output. This averaging process results in a more accurate and stable prediction than relying on any single decision tree.

Handling Complexity: Random Forests are powerful because they can handle a large number of features and complex relationships between features and the target variable. Each tree, being a part of the forest, contributes its understanding of these relationships. When combined, they provide a comprehensive view of how different weather variables affect energy consumption.

4. RESULTS AND DISCUSSION

4.1 Implementation Description

This research implements a regression analysis to explore the relationship between weather patterns and energy consumption in smart homes. Below is the step-by-step explanation:

1. Importing Libraries: The project begins by importing the necessary libraries for data manipulation, analysis, and machine learning. These libraries include pandas for handling data, numpy for numerical operations, matplotlib and seaborn for data visualization, and machine learning libraries from sklearn for regression analysis.

2. Loading the Dataset: The dataset is loaded using pd.read_csv(), assuming the dataset is stored in a CSV file named "Data.csv." The dataset is stored in the df DataFrame.

3. Data Processing

—Handling Missing Data: The code checks for missing values in the dataset using df.isnull().sum().Missing values are dropped from the dataset using df.dropna(inplace=True).

—Exploring the Dataset: The code provides an overview of the dataset using df.info(), df.head(), and df.describe() to check the data's structure, the first few rows, and basic statistics.

—Data Visualization: There is some data visualization using matplotlib and seaborn. For example, it creates a histogram of the "total load forecast" column using plt.hist().

—Correlation Analysis: The code calculates the correlation matrix between numerical variables in the dataset using df.corr(). It also displays correlations in descending order with respect to the "total load forecast" variable.

4. Data Splitting: The dataset is split into training and testing sets using train_test_split from sklearn. The features (X) and the target variable (y) are separated, and standard scaling is applied to the features using StandardScaler.

5. Model Training

—Decision Tree Regressor: A Decision Tree Regressor is created and trained using the training data with hyperparameters like max depth, min samples split, and min samples leaf specified.

—Random Forest Regressor: A Random Forest Regressor is created and trained using the training data.

6. Model Evaluation: The code evaluates the performance of the Decision Tree and Random Forest regressors using the coefficient of determination (R-squared) as a metric, which measures the goodness of fit of the models to the data.

4.2 Results and description

Figure 3 depicts the graphical user interface (GUI) of Smart Homes, presumably showcasing the interface through which users interact with various smart home functionalities. This interface includes features such as controlling lighting, thermostats, security systems, and other connected devices remotely or through automated schedules. The GUI could also display real-time data about energy consumption, indoor air quality, or other relevant metrics, allowing users to make informed decisions about managing their home environment. Figure 4, it presents the data preprocessing steps undertaken before applying a linear regression model. Data preprocessing is a crucial step in machine learning pipelines as it involves cleaning, transforming, and organizing raw data to make it suitable for analysis. This figure includes processes such as handling missing values, scaling features, encoding categorical variables, and splitting the data into training and testing sets. Additionally, the R2 score of the linear regression model is displayed, which indicates the proportion of the variance in the dependent variable that is predictable from the independent variables.

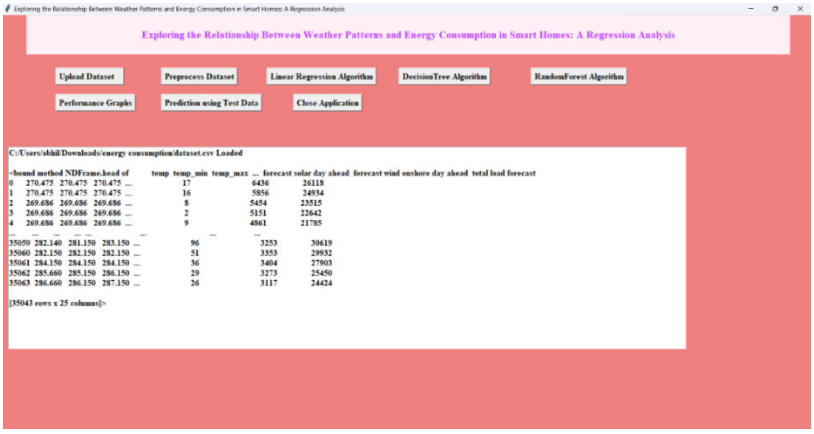

Figure 3: Presents the GUI of Smart Homes.

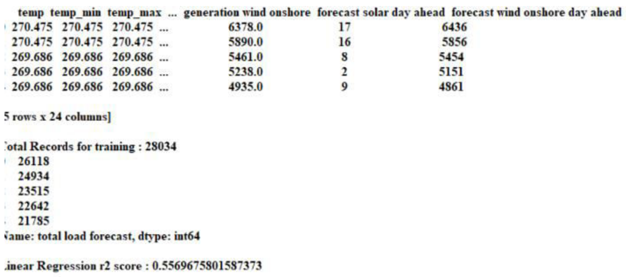

Figure 4: Presents the Data Preprocessing and R2 Score of Linear Regression model.

Figure 5 shows the results of applying the linear regression model to the test data. This includes a plot comparing the actual values of the dependent variable against the predicted values generated by the model. The accuracy of the model's predictions can be evaluated visually by observing how closely the predicted values align with the actual values. Moving on to Figure 6, it displays the R2 score of a decision tree regression model. Decision tree regression is a non-parametric supervised learning method used for regression tasks. The R2 score provides insight into how well the decision tree model fits the data, with a score closer to 1 indicating a better fit.

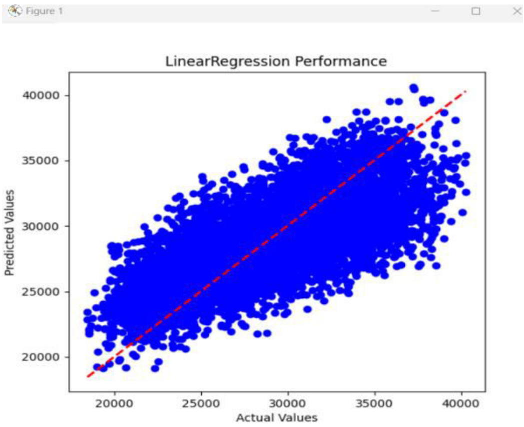

Figure 5: Presents the Linear Regression Model Prediction on Test data.

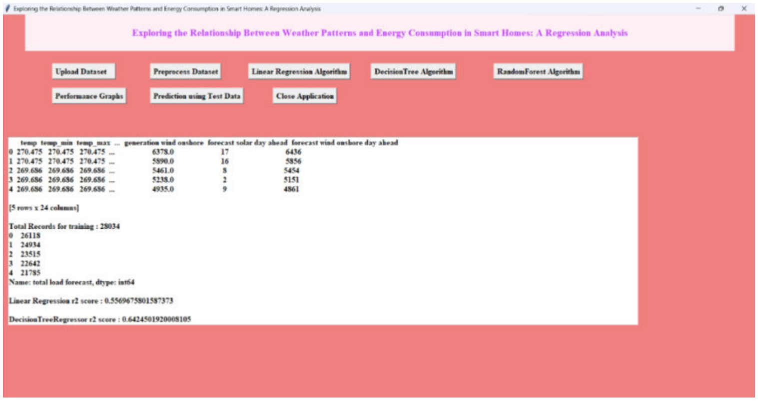

Figure 6: Presents the R2 Score of Decision Tree Regression model.

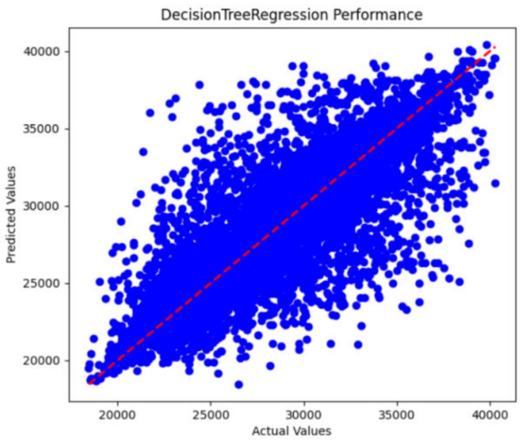

Figure 7: Presents Plot of Decision Tree Regression Model Prediction on Test data.

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876
Research Paper Journal Volume 13, Iss 4, 2023

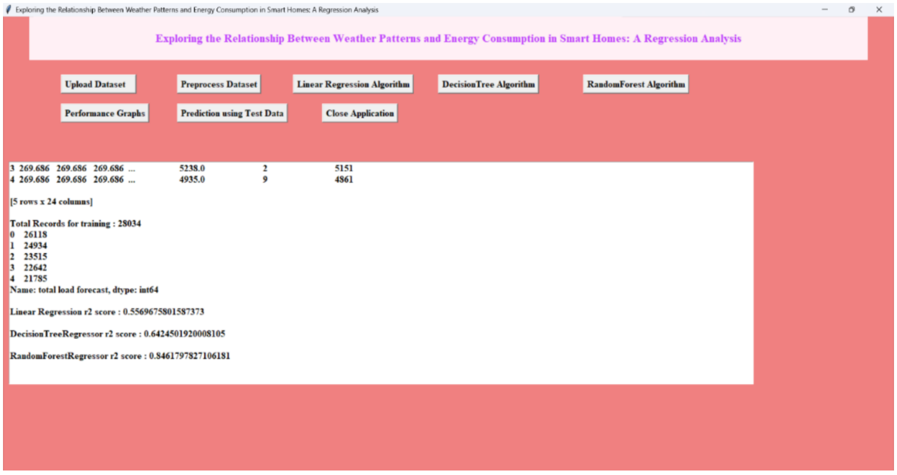

Figure 8: Presents the R2 Score of Random Forest Regression model.

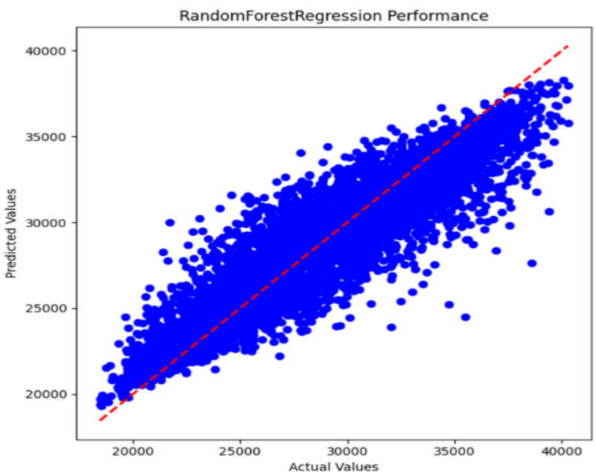

Figure 9: Presents Plot of Random Forest Regression Model Prediction on Test data.

Figure 7 presents a plot illustrating the predictions made by the decision tree regression model on the test data. Similar to Figure 3, this plot compares the actual values against the predicted values generated by the decision tree model. Figure 8 showcases the R2 score of a random forest regression model. Random forest is an ensemble learning method that constructs multiple decision trees during training and outputs the average prediction of the individual trees. The R2 score in this figure provides an indication of the random forest model's predictive performance. Figure 9 displays a plot depicting the predictions made by the random forest regression model on the test data. As with Figures 3 and 5, this plot compares the actual values against the predicted values generated by the random forest model. Figure 10 presents a comparison graph illustrating the performance of each model (linear regression, decision tree regression, and random forest regression) based on their respective R2 scores. This graph allows for a direct comparison of how well each model fits the data and makes predictions. Finally, Figure 11 showcases the predictions made by the proposed model on the test data. This represent a novel or improved regression model developed for a specific application, with its performance evaluated against existing models through metrics such as R2 score or visual comparison of predicted versus actual values.

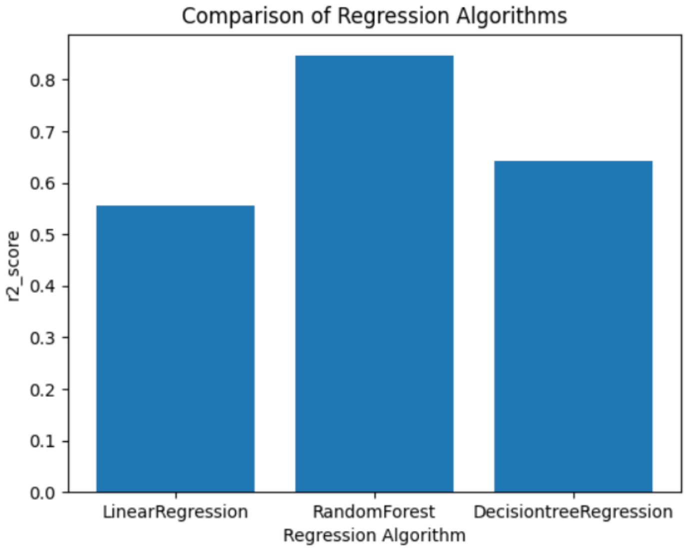

Figure 10: Comparison Graph of each model Performance.

```
Row 1: [2.70475e+02 2.70475e+02 2.70475e+02 1.00100e+03 7.70000e+01 1.00000e+00
 6.20000e+01 0.00000e+00 0.00000e+00 0.00000e+00 0.00000e+00 8.00000e+02
 1.05100e+03 1.89900e+03 0.00000e+00 7.09600e+03 4.30000e+01 7.30000e+01
 4.90000e+01 1.96000e+02 0.00000e+00 6.37800e+03 1.70000e+01 6.43600e+03], Predicted = 14232840.10313622S Watts

Row 2: [2.70475e+02 2.70475e+02 2.70475e+02 1.00100e+03 7.70000e+01 1.00000e+00
 6.20000e+01 0.00000e+00 0.00000e+00 0.00000e+00 0.00000e+00 8.00000e+02
 1.00900e+03 1.65800e+03 0.00000e+00 7.09600e+03 4.30000e+01 7.10000e+01
 5.00000e+01 1.95000e+02 0.00000e+00 5.89000e+03 1.60000e+01 5.85600e+03], Predicted = 12973065.096900636 Watts

Row 3: [ 269.686  269.686  269.686 1002.     78.      0.     23.      0.
   0.      0.      0.    800.    973.   1371.      0.   7099.
  43.     73.     50.    196.      0.   5461.      8.   5454.  ], Predicted = 11538300.223302031 Watts

Row 4: [2.69686e+02 2.69686e+02 2.69686e+02 1.00200e+03 7.80000e+01 0.00000e+00
 2.30000e+01 0.00000e+00 0.00000e+00 0.00000e+00 0.00000e+00 8.00000e+02
 9.49000e+02 7.79000e+02 0.00000e+00 7.09800e+03 4.30000e+01 7.50000e+01
 5.00000e+01 1.91000e+02 0.00000e+00 5.23800e+03 2.00000e+00 5.15100e+03], Predicted = 9333888.95026676 Watts

Row 5: [ 269.686  269.686  269.686 1002.     78.      0.     23.      0.
```

Figure 11: Proposed Model Prediction on test data.

5. CONCLUSION

In conclusion, this research has successfully delved into the intricate relationship between weather patterns and energy consumption in smart homes, employing advanced regression analysis and machine learning techniques. Through meticulous data analysis, meaningful patterns have been extracted, shedding light on the impact of weather variables such as temperature, humidity, and precipitation on energy load. The developed regression models, particularly the decision tree and random forest algorithms, have showcased promising accuracy in predicting energy consumption under varying weather conditions. These findings hold substantial implications for homeowners, energy providers, and policymakers alike. For homeowners, this study provides actionable insights into optimizing energy usage based on weather forecasts. By understanding how weather influences energy consumption, homeowners can implement targeted strategies to reduce costs and enhance efficiency. Energy providers can benefit from these insights by improving demand forecasting and management, ensuring a stable and efficient energy supply. Policymakers can integrate these findings into energy policies, fostering sustainable practices and guiding urban planning initiatives. Furthermore, this work demonstrates the power of data analytics and machine learning in addressing real-world challenges, showcasing their potential in the realm of energy management and sustainability.

REFERENCES

[1] Ribeiro Serrenh, T, and Bertoldi, P. (2019). Smart Home and Appliances: State of the art. Luxembourg: Publications Office of the European Union. doi: 10.2760/453301

[2] Yildiz, B., Bilbao, J. I., Dore, J., and Sproul, A. B. (2017). Recent advances in the analysis of residential electricity consumption and applications of smart meter data. Appl. Energy 208, 402–427. doi: 10.1016/j.apenergy.2017.10.014

[3] Sovacool, B. K., Hook, A., Sareen, S., and Geels, F. W. (2021). Global sustainability, innovation and governance dynamics of national smart electricity meter transitions. Glob. Environ. Change 68:102272. doi: 10.1016/j.gloenvcha.2021.102272

[4] U.S. Energy Information Administration (EIA) (2023). Available online at: https://www.eia.gov/index.php (accessed September 26, 2023).

[5] Ku, A. L., Qiu, Y., Lou, J., Nock, D., and Xing, B. (2022). Changes in hourly electricity consumption under COVID mandates: a glance to future hourly residential power consumption pattern with remote work in Arizona. Appl. Energy 310:118539. doi: 10.1016/j.apenergy.2022.118539

[6] Chinthavali, S., Tansakul, V., Lee, S., Whitehead, M., Tabassum, A., Bhandari, M., et al. (2022). COVID-19 pandemic ramifications on residential smart homes energy use load profiles. Energy Build. 259:111847. doi: 10.1016/j.enbuild.2022.111847

[7] Raman, G., and Peng, J. C.-H. (2021). Electricity consumption of Singaporean households reveals proactive community response to COVID-19 progression. Proc. Natl. Acad. Sci. 118:e2026596118. doi: 10.1073/pnas.2026596118

[8] Li, L., Meinrenken, C. J., Modi, V., and Culligan, P. J. (2021). Impacts of COVID-19 related stay-at-home restrictions on residential electricity use and implications for future grid stability. Energy Build. 251:111330. doi: 10.1016/j.enbuild.2021.111330

[9] Lou, J., Qiu, Y., Ku, A. L., Nock, D., and Xing, B. (2021). Inequitable and heterogeneous impacts on electricity consumption from COVID-19 mitigation measures. iScience 24:103231. doi: 10.1016/j.isci.2021.103231

[10] Sánchez-López, M., Moreno, R., Alvarado, D., Suazo-Martínez, C., Negrete-Pincetic, M., Olivares, D., et al. (2022). The diverse impacts of COVID-19 on electricity demand: the case of Chile. Int. J. Electr. Power Energy Syst. 138:107883. doi: 10.1016/j.ijepes.2021.107883

[11] Qiu, Y. L., Wang, Y. D., Iseki, H., Shen, X., Xing, B., and Zhang, H. (2022a). Empirical grid impact of in-home electric vehicle charging differs from predictions. Resour. Energy Econ. 67:101275. doi: 10.1016/j.reseneeco.2021.101275

[12] Al Khafaf, N., Rezaei, A. A., Moradi Amani, A., Jalili, M., McGrath, B., Meegahapola, L., et al. (2022). Impact of battery storage on residential energy consumption: an Australian case study based on smart meter data. Renew. Energy 182, 390–400. doi: 10.1016/j.renene.2021.10.005

[13] Qiu, Y., Xing, B., Patwardhan, A., Hultman, N., and Zhang, H. (2022b). Heterogeneous changes in electricity consumption patterns of residential distributed solar consumers due to battery storage adoption. iScience 25:104352. doi: 10.1016/j.isci.2022.104352

[14] Liang, J., Qiu, Y., and Xing, B. (2022a). Impacts of electric-driven heat pumps on residential electricity consumption: an empirical analysis from Arizona, USA. Clean. Responsible Consum. 4:100045. doi: 10.1016/j.clrc.2021.100045