# Artificial Intelligence Project

Submitted By-

Naman Bansal: 102303084

Prithvi: 102303090

Submitted To-

Dr. Simranjit Kaur



Index

1. Abstract

2. Introduction

3. Problem Statement

4. Methodology

5. Tools and Technologies

6. Expected Outcomes

7. Pre-processing and cleaning of DataSet

8. Model comparison for building Energy Usage Prediction System

9. Prediction of Energy of Different Appliances

10. Conclusion

11. References

# Abstract

# 

With the increasing adoption of smart home technologies, optimizing energy consumption has become a critical aspect of sustainable living. This project aims to develop an intelligent and accurate energy usage prediction system for household appliances using machine learning techniques. By leveraging sensor data and temporal features such as hour and weekday, the system forecasts appliance-specific energy consumption to support efficient energy management. A comprehensive data preprocessing pipeline was employed, including missing value treatment, feature engineering, and encoding of categorical data. Various regression algorithms—including Linear Regression, Decision Tree, Support Vector Regression, Random Forest, and Gradient Boosting—were trained and evaluated on multiple performance metrics such as R² score, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). After comparative analysis, Gradient Boosting and Random Forest emerged as the top-performing models due to their strong predictive accuracy and robustness. Additionally, a multi-output regression model was implemented to simultaneously predict energy usage across multiple appliances. The resulting system highlights the practical application of machine learning in enhancing energy efficiency in smart homes, offering actionable insights for reducing consumption, lowering costs, and promoting environmentally responsible behavior.

## Introduction

Project Title: Energy Prediction and Optimization in Smart Homes Using Machine Learning

In today's era of growing energy demands and environmental concerns, smart homes have emerged as a promising solution for sustainable living. With the integration of sensors, smart devices, and data analytics, smart homes offer opportunities to monitor and optimize energy consumption in real-time. However, efficient energy management still remains a challenge due to varying usage patterns, unpredictable weather conditions, and appliance behavior.

This project focuses on the application of machine learning (ML) techniques to predict and optimize energy usage within a smart home environment. Traditional energy-saving approaches often rely on static rules or manual configurations, which fail to adapt dynamically to changing conditions. By leveraging historical energy data, ambient conditions (temperature, humidity, visibility, etc.), and appliance usage logs, our system aims to provide intelligent recommendations to reduce energy waste and enhance efficiency.

The project incorporates supervised learning models for energy consumption prediction, anomaly detection for identifying unusual power usage (e.g., malfunctioning appliances), and appliance control strategies based on predicted load. In addition, smart scheduling is proposed to run high-load appliances during off-peak hours, and integration with solar generation data is explored to promote renewable usage.

### **Problem Statement**

To develop a machine learning-based system that accurately predicts appliance-level energy consumption in a smart home environment using historical sensor and temporal data, with the goal of optimizing energy usage, reducing operational costs, and supporting sustainable living.

## Methodology

## Data Collection & Preprocessing:

The dataset includes real-time smart home energy consumption data along with environmental factors such as temperature, humidity, visibility, pressure, wind speed, and cloud cover. Data cleaning is performed to handle missing values, convert datetime fields, and normalize numerical features for consistent scaling.

Feature Engineering:

Extract relevant temporal features (e.g., hour of day, day of week), encode categorical variables, and generate lag features or rolling averages to capture energy usage trends over time. Correlation analysis is used to select impactful predictors.

Model Selection & Training:

Multiple machine learning algorithms such as Linear Regression, Random Forest Regressor, K-Nearest Neighbors, and Gradient Boosting are trained to predict total energy consumption. For classification-based tasks like appliance state prediction and anomaly detection, models like Logistic Regression, Decision Trees, and Isolation Forest are employed.

Optimization Strategies:

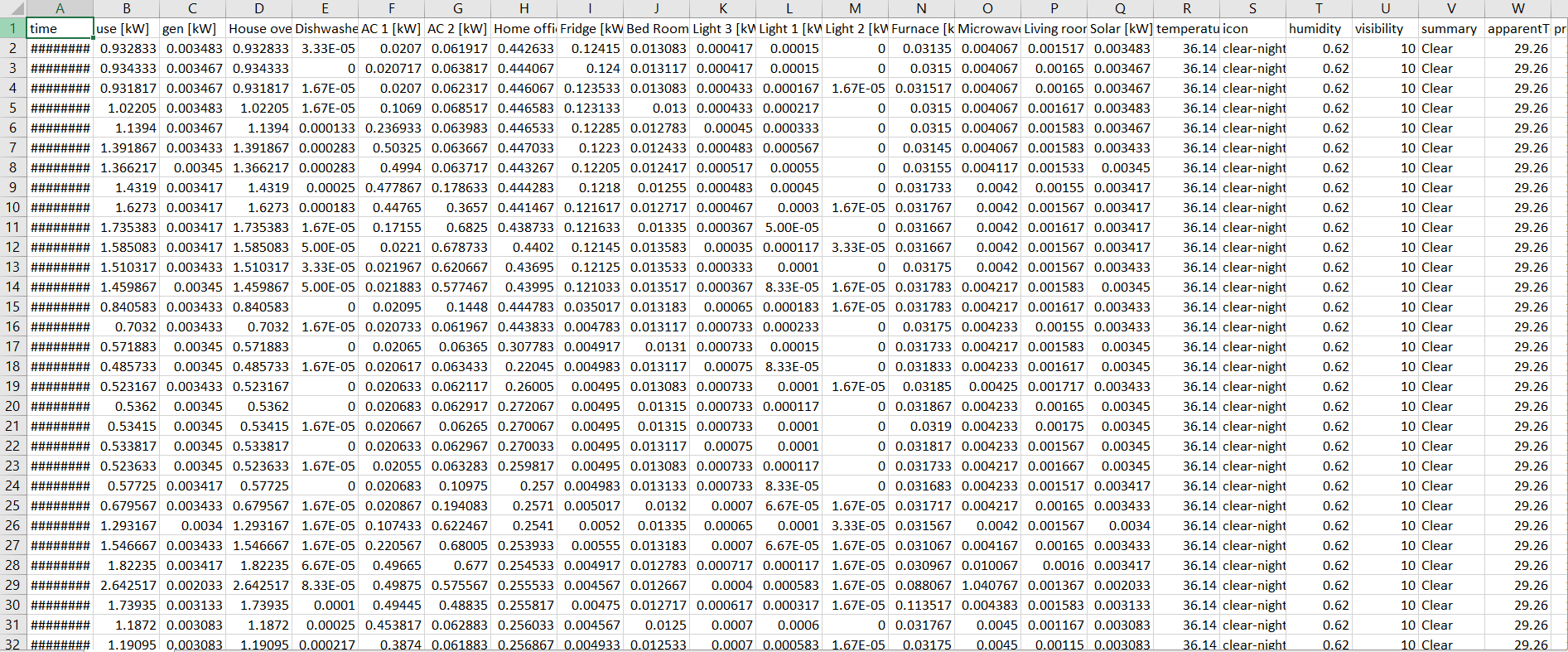
Predictive insights are used to schedule heavy-load appliances (e.g., furnace, AC, fridge) during low energy usage hours. The system can also suggest shutting off certain appliances based on predicted peak consumption and weather conditions. Solar generation forecasts are used to align consumption with renewable availability.

Evaluation:

Models are assessed using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R² Score for regression tasks, and Accuracy, Precision, Recall, and F1-Score for classification/anomaly detection. Confusion matrices and residual plots are also generated.

### **Tools & Technologies**

**• Python, Jupyter Notebook** – Primary programming language and development environment  
**• Pandas, NumPy** – For data manipulation, preprocessing, and numerical computations  
**• Scikit-learn** – For implementing and evaluating machine learning regression models  
**• Machine Learning Algorithms:**  
  • Linear Regression  
  • Decision Tree Regressor  
  • Random Forest Regressor  
  • Support Vector Regressor (SVR)  
  • Gradient Boosting Regressor  
**• MultiOutputRegressor** – For multi-appliance prediction  
**• Data Visualization: Matplotlib,** **Seaborn** – Used to explore data distributions, trends, and model performance  
**• Dataset**: <https://www.kaggle.com/datasets/taranvee/smart-home-dataset-with-weather-information>



Expected Outcomes:

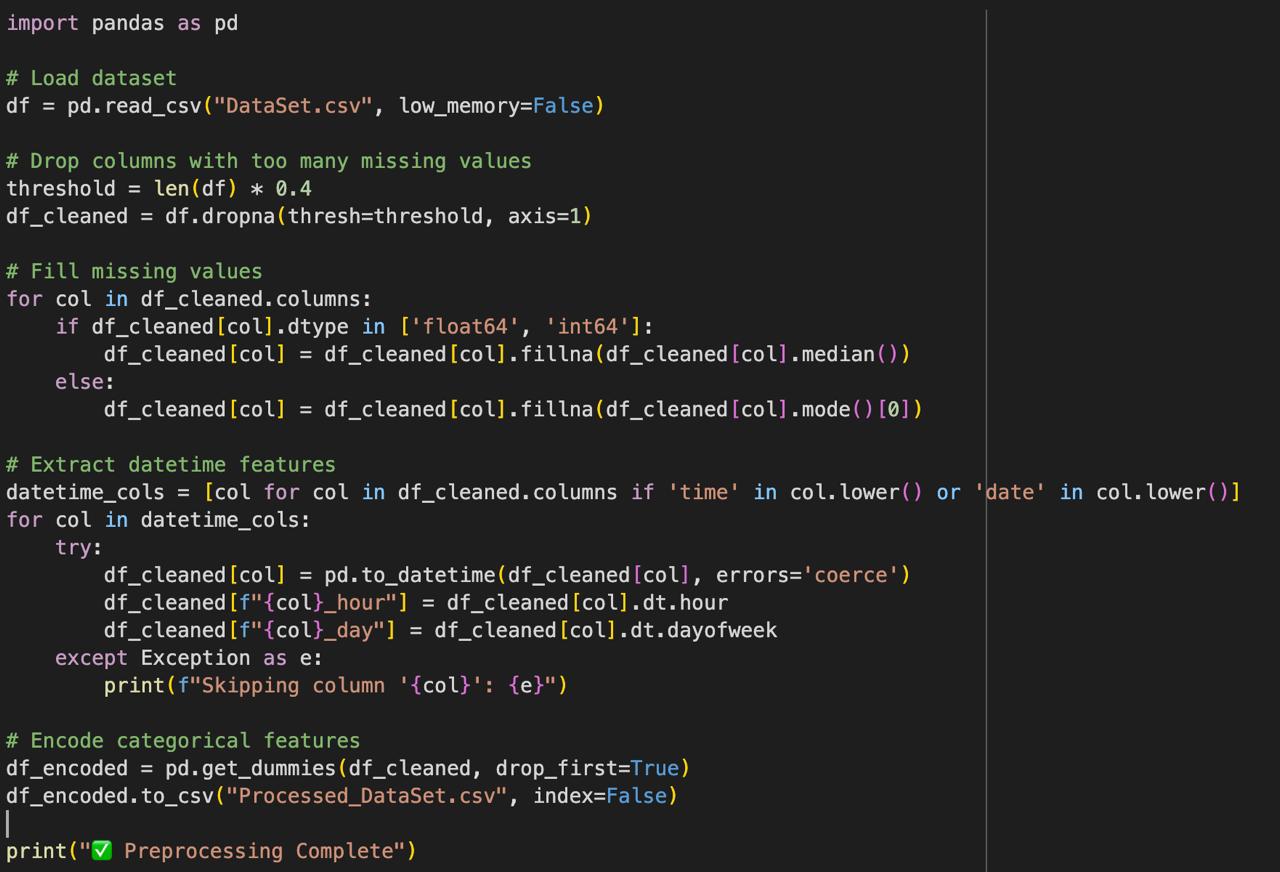
A functional machine learning-based system capable of accurately predicting energy consumption in a smart home environment.

Improved energy efficiency through intelligent scheduling and optimization of high-load appliances based on usage patterns and environmental conditions.

Integration of renewable energy insights (e.g., solar generation) to promote sustainable energy usage.

Increased awareness of the role of artificial intelligence and machine learning in modern home automation and energy management systems.

Pre-Processing and Cleaning of DataSet



### **Model Comparison for Building Energy Usage Prediction System**

To build a reliable and accurate energy prediction system for smart home appliances, multiple machine learning regression models were implemented and compared. These models were evaluated based on their ability to predict appliance-level energy consumption using sensor and temporal data. The models include:

• **Linear Regression**: A baseline model used to understand linear relationships between input features and appliance energy usage.

• **Decision Tree Regressor**: A non-linear model that creates a tree-based structure to make predictions, useful for interpreting feature importance and splits.

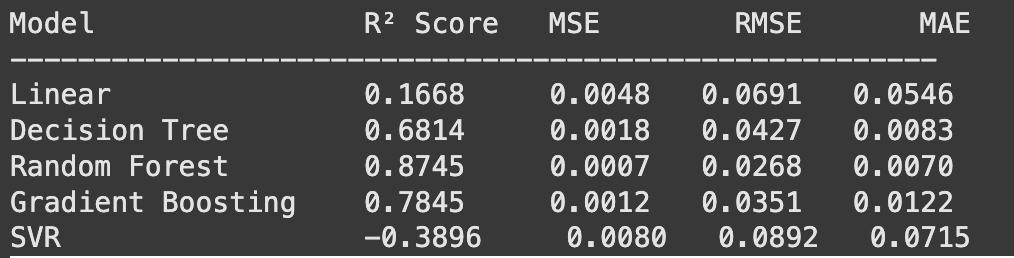
• **Support Vector Regressor (SVR)**: Effective for high-dimensional regression problems, though computationally intensive with large datasets.

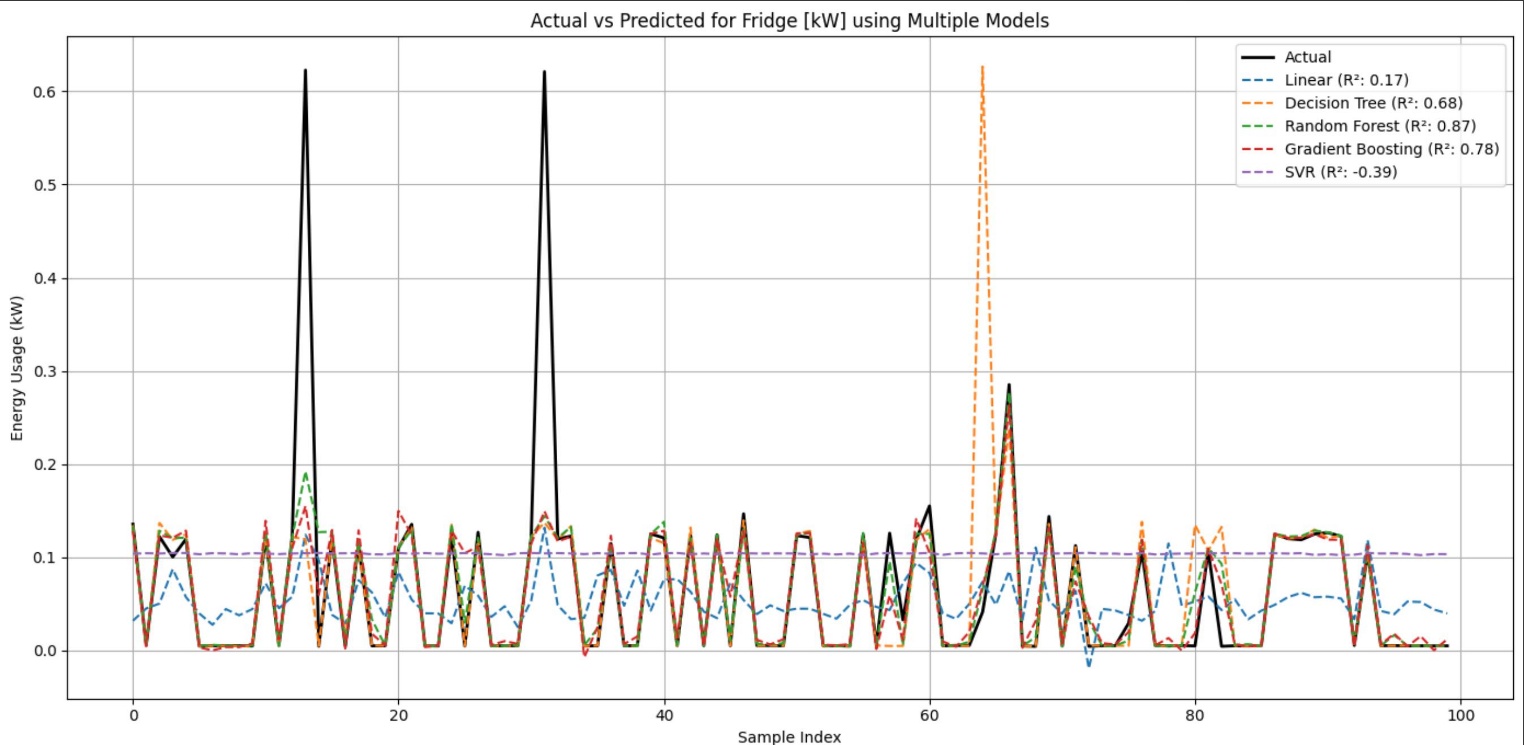
• **Random Forest Regressor**: An ensemble of decision trees that improves accuracy and reduces overfitting through bootstrapping and averaging techniques.

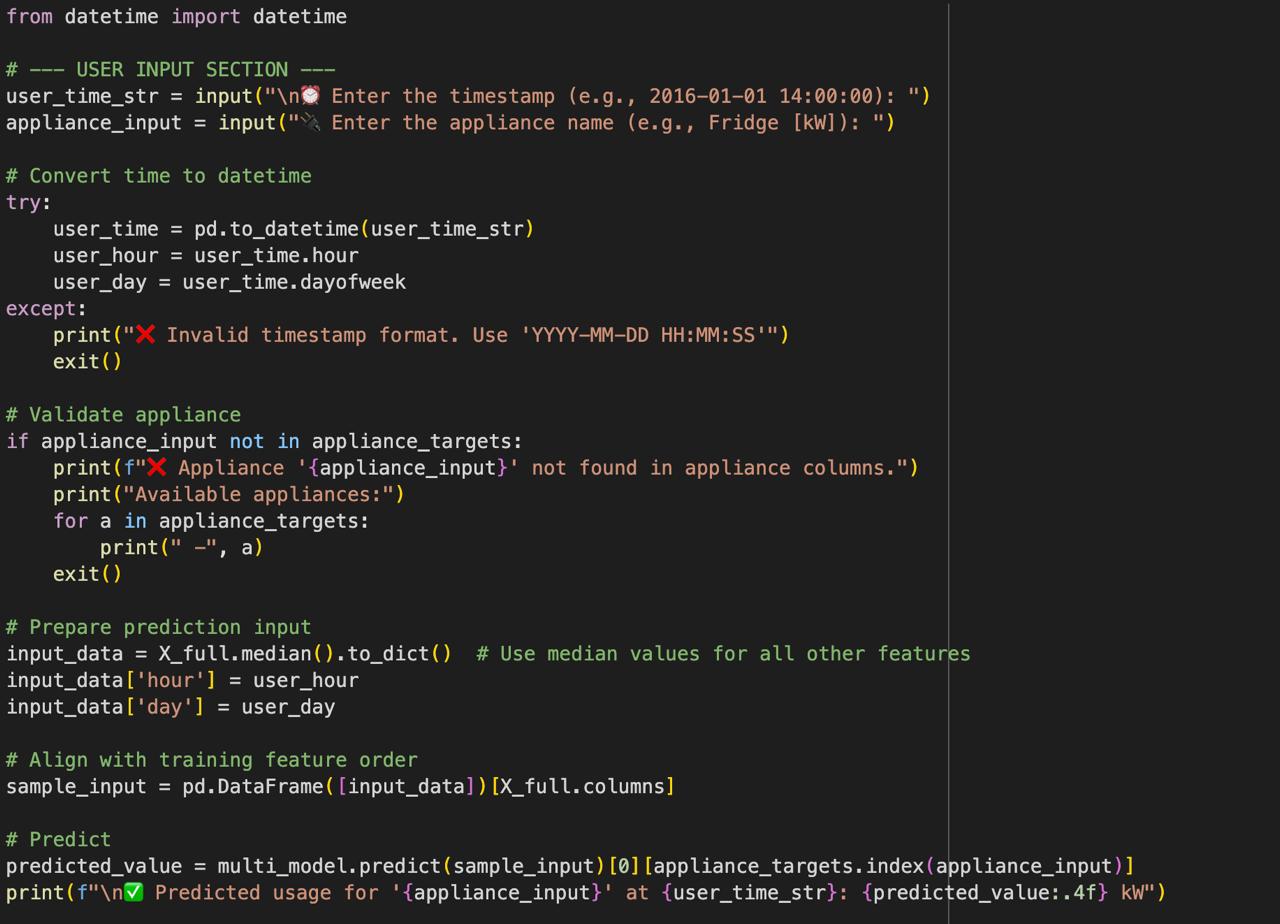
• **Gradient Boosting Regressor**: Builds models sequentially by minimizing residual errors, delivering high accuracy and robustness in predictive tasks.

• **MultiOutputRegressor**: Used to extend the best-performing model to predict energy consumption across multiple appliances simultaneously.

Each model was evaluated using metrics such as **R² score**, **Mean Absolute Error (MAE)**, and **Root Mean Squared Error (RMSE)**. Random Forest and Gradient Boosting Regressors consistently outperformed others in terms of both accuracy and generalization, making them suitable for deployment in smart energy management systems.







## 

## Prediction of Energy of different appliances

### Conclusion

The proposed smart home energy usage prediction system leverages machine learning techniques to forecast appliance-level power consumption based on historical and sensor-based data. Multiple regression algorithms—including Linear Regression, Decision Tree, Support Vector Regression, Random Forest, and Gradient Boosting—were implemented and evaluated using performance metrics such as R² score, RMSE, and MAE. After thorough experimentation and model comparison, the Gradient Boosting and Random Forest Regressors emerged as the most accurate and robust solutions.

The selected models were further trained using a multi-output regression approach to enable simultaneous prediction of multiple appliance energy usages. This enhances the system’s practicality for real-world applications in smart homes. The predictive system can be integrated into home automation platforms, enabling residents to monitor and optimize energy usage in real time.

This project not only contributes to improving energy efficiency and reducing operational costs but also demonstrates the power of artificial intelligence and data science in building sustainable, intelligent environments. The framework provides a foundation for future developments, such as real-time deployment with IoT integration, adaptive energy control, and scalable energy forecasting for smart communities.

References

* <https://www.kaggle.com/datasets/taranvee/smart-home-dataset-with-weather-information>
* <https://ieeexplore.ieee.org/document/9243578>