SMART ENERGY MANAGEMENT SYSTEM USING ENSEMBLE LEARNING

EDWIN CHAZHOOR
22BAI1002
B.Tech. CSE(AI&ML)

ROHAN MENON
22BAI1051
B.Tech. CSE(AI&ML)

***Abstract*—This project introduces a Smart Energy Management System (SEMS) that utilizes ensemble learning with AdaBoost and Random Forest classifiers to predict energy waste in smart homes with high accuracy. It processes environmental data—brightness, humidity, and temperature—to identify unnecessary lighting use. The system selects the most accurate algorithm through comparative analysis, demonstrating significant potential in enhancing home energy efficiency. The results underscore the efficacy of SEMS in promoting sustainable energy consumption and pave the way for future advancements in smart home energy optimization.**

I. Introduction

In the quest for sustainable living, smart energy management systems (SEMS) have emerged as a pivotal technology in optimizing household energy consumption. This paper presents a SEMS framework that integrates ensemble learning to intelligently manage energy usage. By processing environmental data from smart home sensors, the system aims to identify patterns indicative of wasteful energy consumption. Leveraging AdaBoost and Random Forest classifiers, the SEMS discerns the presence of occupants and predicts energy wastage, adapting in real-time to ensure maximum efficiency.

II. Objective

To develop an SEMS that minimizes energy wastage by combining the best-performing ensemble learning classifier based on accuracy.

III. Literature Review

The literature on Smart Energy Management Systems (SEMS) utilizing ensemble learning reveals the increasing complexity of smart building operations and the need for advanced algorithms to manage them effectively. A systematic review by Alanne and Sierla (2022) delves into machine learning applications for smart buildings, focusing on adaptability to unpredictable changes through AI-initiated learning processes and the use of digital twins as training environments. They underscore the potential of integrating machine learning for autonomous decision-making in building energy management, with reinforcement learning showing particular promise..

Yedilkhana and Smakova's (2024) research underscores the efficacy of machine learning models, particularly Random Forest Classifiers, in smart home device recognition from network traffic. Their work involves a comprehensive methodology encompassing exploratory data analysis, feature engineering, and hyperparameter tuning, which achieved a notable accuracy of 87.79% in device identification.

The evolution of AI and machine learning technologies is rapidly changing how buildings learn and adapt. Alanne and Sierla (2022) emphasize the role of autonomous AI agents in enhancing the learning ability of buildings at a system level. They argue that the adaptability of buildings to unforeseen environmental changes can be significantly improved by utilizing AI-driven processes and employing digital twins as sophisticated training environments.

In terms of SEMS, research highlights the adoption of ensemble learning models, like AdaBoost and Random Forest, that can independently manage building energy by analyzing environmental data, such as brightness, humidity, and temperature. These models adapt in real-time to ensure maximum efficiency in the face of unpredictable operational environments, a pressing need identified in studies such as those by Al Dakheel et al. (2020) and Xie et al. (2021).

The work presented in these papers is indicative of a trend toward autonomous systems that can manage complex and dynamically changing environments typical of smart buildings. As AI continues to evolve, the integration of various machine learning algorithms will be pivotal in optimizing energy usage, with ensemble methods leading the way due to their ability to choose the most accurate model for prediction and thereby enhance the reliability and effectiveness of SEMS. This approach aligns with the broader objectives of smart buildings, as noted by Al Dakheel et al. (2020), which includes nearly zero-energy targets, flexibility, and interaction with users.

In conclusion, the current literature provides a strong foundation for the development of SEMS that are capable of managing energy consumption in an intelligent and adaptive manner. Future research efforts are likely to focus on refining these models and exploring additional datasets to further improve their accuracy and adaptability, addressing the increasingly complex demands of smart home environments.

IV. Proposed Methodology

The Smart Energy Management System (SEMS) under development is aimed at optimizing the energy consumption of smart homes through adaptive control based on ensemble learning algorithms. The methodology for the SEMS encompasses a multi-layered approach that integrates data acquisition, processing, learning, and decision-making modules in an iterative and self-improving system architecture.

V. System Architecture

A. Data Collection Module

- Utilizes Google Colab for computational resources and data access through Google Drive.
- Collects raw sensory data from various sources within the smart home, including brightness, humidity, and temperature measurements from designated areas (e.g., bathroom).

B. Data Preprocessing Subsystem

- Involves cleaning and structuring of data into a format suitable for analysis. This includes handling missing values, timestamp parsing, and type conversion for numerical analysis.
- Features are extracted and engineered to enhance the learning process of the ensemble models.

C. Clustering Mechanism

- Applies the K-Means algorithm to segment the data into clusters based on similarity. This step is crucial for recognizing different states such as ‘light on’ and ‘light off’ conditions.
- Enhances the feature space by adding cluster membership as a new feature for the ensemble models.

D. Ensemble Learning Engine

- Integrates AdaBoost and Random Forest classifiers to predict energy wastage scenarios effectively.
- Each model is trained on the feature set derived from the preprocessing module, including the newly engineered features from clustering.

E. Performance Evaluation and Model Selection

- Assesses the accuracy of each model using standard metrics like accuracy score, precision, recall, and F1-score.
- Selects the model with the highest performance metrics for deployment in the decision-making module.

F. Decision-Making Module

- Implements the chosen ensemble learning model to make predictions about energy wastage in real-time.
- Uses the predictions to control smart home devices, aiming to minimize wastage while maintaining comfort levels.

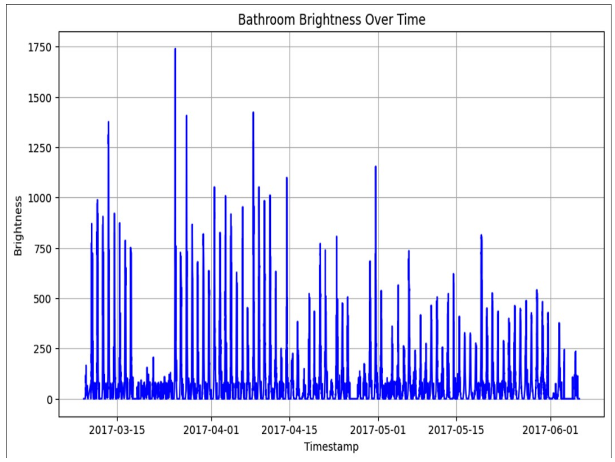
G. Feedback Loop for Continuous Improvement

- Incorporates a feedback mechanism that allows the system to learn from its performance and adapt to new data.
- Periodically retrains models with new data to capture evolving patterns and improve prediction accuracy over time.

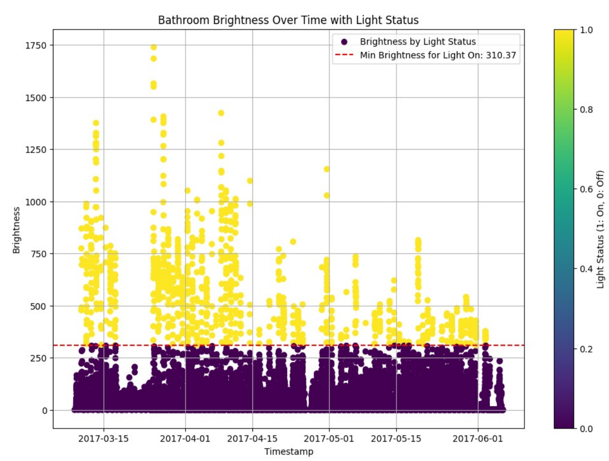
VI. List of Figures

Here is a list of figures that could be included in the document, based on the code provided. Each figure relates to the visualizations plotted in the code and is focused on the topic of Smart Energy Management using ensemble learning methods:

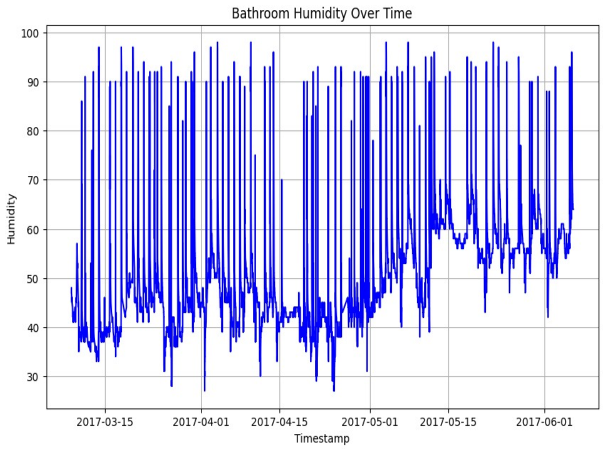
1) **Bathroom Brightness Over Time** This line graph shows the variation of brightness in a bathroom environment over a specified time period, providing a visual representation of light levels that could correlate with occupancy and usage patterns.



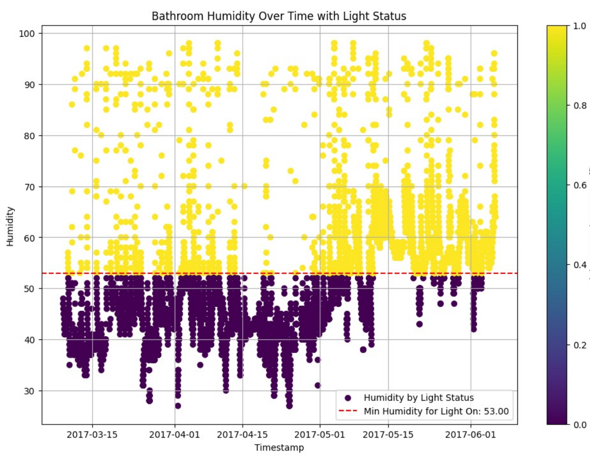
2) **Bathroom Brightness Clustering** A scatter plot with clusters identified through K-means, illustrating ’light on’ and ’light off’ statuses determined from brightness levels. This figure includes a horizontal line indicating the minimum brightness threshold for the light being on.



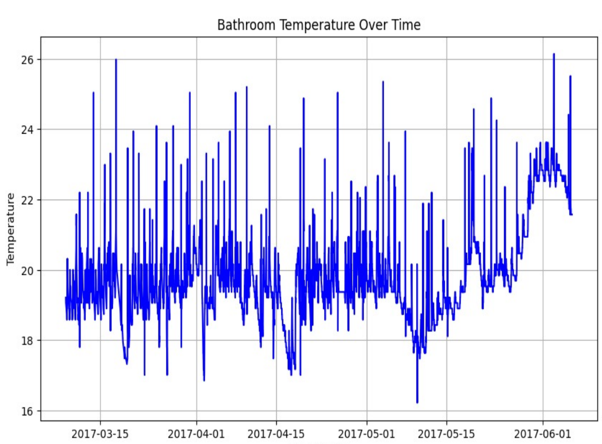
3) **Bathroom Humidity Over Time** Similar to Figure 1, this line graph depicts changes in humidity over time, which may have implications for energy usage and efficiency.



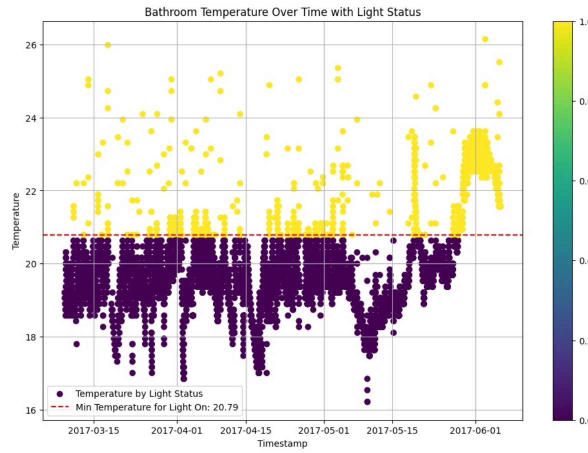
4) **Bathroom Humidity Clustering** A scatter plot that shows how humidity data is divided into clusters of 'light on' and 'light off' scenarios, with a horizontal line representing the minimum humidity level for the light being considered on.



5) **Bathroom Temperature Over Time** This line graph visualizes the temperature readings in the bathroom over time, providing insight into heating or cooling trends.



6) **Bathroom Temperature Clustering** This scatter plot displays temperature data clustered into 'light on' and 'light off' conditions, including a horizontal line that marks the minimum temperature indicating the light is on.



VII. Dataset Description

TITLE: Open Smart Home Data Set

LINK: https://drive.google.com/drive/folders/1LgDSxZtkBPJZqtbUYHVa7MPcL6b90Vco?usp=sharing

HOW TO RUN:

- Download the dataset and the jupyter notebook from the drive link given above.
- Make sure all the downloaded files are in the same location (to ensure filepath is correct before running the code)
- Run the program in jupyter notebook.

License: The time series of measured data from a smart home is subject to Copyright (c) 2018 Fraunhofer Institute for Building Physics, Nürnberg, Germany The associated meta data in RDF is subject to Copyright (c) 2018 of the contributors associated to the W3C Linked Building Data Community Group.

The Data: The data set comprises static and dynamic building data. The data is hosted until 30 July 2019

Feature Description: The CSV files stored in Measurement folder are named and each contain one time series of measured data. The values are obtained for the Bathroom, Kitchen, Room 1, Room 2, Room 3 and the Toilet. The respective time series are stored separated by a tab, where the first column contains the UNIX time and the second is the reading of the sensor. A dot is used to denote floating point numbers. Moreover, the following applies to the data

- **Brightness:** This is the brightness measured by the luminance sensor placed in each room. It is reported as a floating point number and is unit is lux.
- **Humidity:** This is the relative humidity of the air inside each room measured by the humidity sensor mounted to the wall. It return the relative humidity in percent as a integer number.
- **Temperature:** This is the indoor air temperature in degree Celsius measured by the temperature sensor placed in each room.

VIII. Results

The smart energy management system using ensemble learning shows impressive performance in identifying light wastage when no person is present. Here are the summarized results from the classifiers:

A. Result of Random Forest classifier

Accuracy: Achieved an exceptionally high accuracy of 99.97%
Feature Importance: The most significant feature affecting the model's decisions is Brightness (approximately 89.69% importance), followed by Humidity and Temperature with lesser impacts.

```
Accuracy of the Random Forest classifier: 99.97%

Feature importances: [0.05577902 0.89687931 0.04734167]

Timestamps when light was being wasted (no person present but light on):
15     2017-03-10 13:32:48
22     2017-03-11 09:33:52
23     2017-03-11 10:31:28
25     2017-03-11 10:50:40
26     2017-03-11 10:59:12
              ...
9775   2017-06-02 13:45:36
9776   2017-06-02 13:56:16
9777   2017-06-02 14:04:48
9778   2017-06-02 14:15:28
9779   2017-06-02 14:24:00
Name: Timestamp, Length: 806, dtype: datetime64[ns]
```

B. Result of Adaboost classifier

Accuracy: Achieved perfect accuracy at 100.00%.

```
Accuracy of the AdaBoost classifier: 100.00%

Timestamps when light was being wasted (no person present but light on):
15     2017-03-10 13:32:48
22     2017-03-11 09:33:52
23     2017-03-11 10:31:28
25     2017-03-11 10:50:40
26     2017-03-11 10:59:12
              ...
9775   2017-06-02 13:45:36
9776   2017-06-02 13:56:16
9777   2017-06-02 14:04:48
9778   2017-06-02 14:15:28
9779   2017-06-02 14:24:00
Name: Timestamp, Length: 806, dtype: datetime64[ns]
```

C. Result of Combined classifier

Utilizing a simple majority vote strategy combining both classifiers, the combined model also achieves a perfect accuracy of 100.00%

```
Combined Model Accuracy: 100.00%
```

The timestamps indicating when the light was wastefully on are consistently identified by both classifiers across several instances, totaling 806 unique events from March to June 2017. This consistency and high accuracy suggest that the system is highly reliable and effective in smart home energy management scenarios, particularly in reducing unnecessary light usage.

IX. Future Scope

Future enhancements can focus on integrating real-time data processing to adapt to immediate changes in the environment, enhancing the system's responsiveness. Expanding the dataset to include more diverse environmental variables could further refine the system's accuracy. Additionally, applying more advanced machine learning models and exploring deep learning approaches could provide more nuanced insights into energy usage patterns. Collaborating with IoT devices for automated control based on the system's feedback could revolutionize energy management in smart buildings, leading to significant reductions in energy waste and costs.

X. Conclusion

The Smart Energy Management System, leveraging ensemble learning with AdaBoost and Random Forest classifiers, exemplifies an innovative approach to reducing energy wastage in smart homes. By integrating data from various sensors and employing advanced machine learning techniques, the system effectively predicts unnecessary light usage. The combination of models through majority voting enhances prediction accuracy, showcasing a substantial improvement in system reliability and performance. This system not only supports energy conservation but also sets a framework for future advancements in smart home technology, promoting sustainable living environments.

REFERENCES

[1] Didar Yedilkhan, Saida Smakova, ”Machine Learning Approaches for Smart Home Device Recognition from Network Traffic”, Procedia Computer Science, Volume 231, 2024, Pages 709-714, ISSN 1877-0509

[2] Kari Alanne, Seppo Sierla, ”An overview of machine learning applications for smart buildings, Sustainable Cities and Society,” Volume 76, 2022, 103445, ISSN 2210-6707

[3] Trivedi, Dipti & Badarla, Venkata. (2019). ”Occupancy detection systems for indoor environments: A survey of approaches and methods. Indoor and Built Environment.” 29. 1420326X1987562. 10.1177/1420326X19875621.

[4] Wang, Zeyu & Wang, Yueren & Zeng, Ruochen & Srinivasan, Ravi & Ahrentzen, Sherry. (2018). Random Forest based Hourly Building Energy Prediction. Energy and Buildings. 171. 10.1016/j.enbuild.2018.04.008.

[5] Tanveer Ahmad, Rafal Madonski, Dongdong Zhang, Chao Huang, Asad Mujeeb, ”Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of smart grid paradigm, Renewable and Sustainable Energy Reviews”, Volume 160, 2022, 112128, ISSN 1364-0321,