Enhancing Activity Monitoring in Smart Homes with IoT enabled Sensor Networks using Machine Learning

Sristi Laxmi Lalitha[1], K. Srujana[2], B. Harshith[2], V. Soumith Reddy[2], B. Akhil[2]

[1]Associate Professor, [2]UG Student, [1,2]Department of CSE (Data Science)
[1,2]Vaagdevi College of Engineering, Bollikunta, Warangal, Telangana, India
Corresponding Email: lalithayagnam03@gmail.com

ABSTRACT

The rapid advancement of the Internet of Things (IoT) has enabled the incorporation of intelligent sensor networks in smart homes, enhancing activity monitoring, security, and automation capabilities. This design emphasizes the creation of an advanced IoT-enabled detector network aimed at real-time exertion monitoring within smart homes, thereby improving security, energy efficiency, and supported living operations. Traditional monitoring systems frequently encounter extended durations of inactivity, constrained data accessibility, and restricted data processing functionalities. This design implements machine learning algorithms to attain accurate human activity recognition and incorporates edge computing to improve the efficiency of real-time data processing. The system will utilize a network of intelligent detectors, including stir sensors, environmental detectors, and wearable devices, to collect and analyze exertion patterns for the purpose of identifying anomalies and automating responses. Advanced machine learning models, including deep learning-based sequence classifiers and anomaly detection algorithms, are expected to enhance the accuracy of activity recognition and security monitoring. Furthermore, the implementation of energy-efficient adaptive literacy methods will lead to a decrease in computational output, which will subsequently improve the scalability and sustainability of the system. Sequestration will entail the application of conservation mechanisms, including secure encryption and authentication protocols, to safeguard sensitive user data. The proposed frame improves home security, facilitates independent living for seniors, and enhances smart home automation, thereby contributing to the creation of safer, more intelligent, and energy-efficient living environments.

Keywords: Activity Monitoring, Sensor Fusion, Human Activity Recognition, Edge Computing in IoT.

1. INTRODUCTION

The integration of IoT-enabled detector networks in smart homes has changed the way individuals interact with their living environments. Smart homes employ interconnected devices and sensors to automate routine activities, bolster security measures, and optimize energy efficiency. The function of exertion monitoring within IoT-enabled smart homes is critical. This process involves the tracking and analysis of human behavior to enhance safety, optimize resource usage, and facilitate personalized automation. Traditional exertion covering systems depend on cameras, stir detectors, and tailored interventions, which may demonstrate constraints in terms of sequestration, real-time processing, and inflexibility. The growing implementation of wireless sensor networks (WSNs), edge computing, and AI-driven analytics facilitates the creation of effective, real-time, and adaptive monitoring solutions that can detect and analyze human behavior with improved accuracy. This design aims to enhance exertion monitoring in smart homes by implementing a network of IoT-enabled detectors, which are integrated with machine learning algorithms. The proposed system will employ advanced smart detectors, which include stir sensors, temperature detectors, wearables, and environmental detectors. These components will be used to collect and analyze data related to various activities, such as walking, sleeping, cooking, exercising, and detecting anomalies, including falls or unusual inactivity. The system utilizes real-time processing techniques to provide immediate alerts and facilitate adaptive automation, which improves the overall efficiency of smart homes.

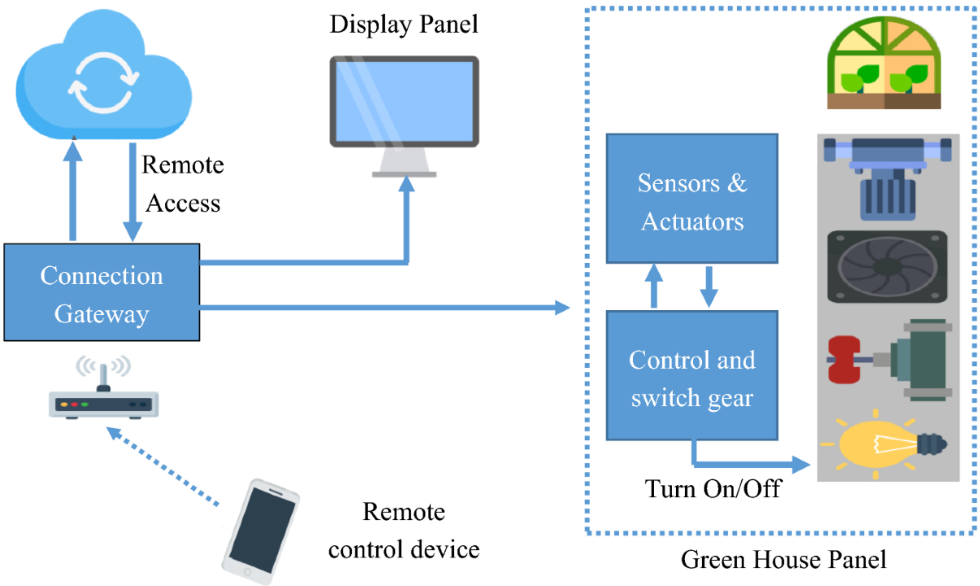

Fig. 1: Overview of IoT architecture in smart homes.

The design also tackles essential challenges including energy efficiency, data sequestration, and scalability. The deployment of edge computing will reduce latency and computational overhead, facilitating real-time processing near the data source. Advanced encryption and security mechanisms are implemented to safeguard sensitive data, thereby ensuring the system's reliability and focus on privacy.

2. LITERATURE SURVEY

The incorporation of IoT-enabled detector networks within smart homes has transformed the manner in which individuals engage with their living environments. Smart homes utilize interconnected devices and sensors to automate routine tasks, improve security protocols, and maximize energy efficiency. The assessment of exertion represents a fundamental capability of Internet of Things (IoT) technology within smart home environments. This procedure encompasses the monitoring and examination of human behavior to guarantee safety, enhance resource efficiency, and enable tailored automation. The analysis of high-frequency electricity data enables the examination of electricity consumption patterns among different consumer groups over specific time intervals, along with the assessment of behavioral changes that occur after the implementation of new technologies and demand-side operational strategies. Furthermore, high-frequency data improves the accuracy of energy consumption forecasts as it exhibits greater variability. The utilization of high-frequency electricity data during epidemic periods has enabled the analysis and assessment of the overall effects of COVID-19 on energy consumption and transition in both pre- and post-pandemic scenarios. The global population has undergone alterations in behavioral patterns and daily routines due to the epidemic. As a result, the patterns of electricity consumption in residential and commercial buildings have changed. Ku et al. [5] employed individual hourly power consumption data within a machine learning framework to examine changes in electricity usage patterns due to COVID-19 regulations in Arizona. Chinthavali et al. [6] conducted an analysis of variations in energy consumption patterns on weekdays and weekends, comparing data from before and after the COVID-19 pandemic. Raman and Peng [7] conducted an analysis of domestic electricity consumption data, revealing a significant positive correlation between the progression of the epidemic and domestic electricity consumption in Singapore. Li et al. conducted an analysis of data from New York apartments to examine the relationship between the number of COVID-19 cases, outdoor temperature, and residential electricity consumption.

Lou et al [8]. determined that the measures implemented in response to COVID-19 led to a 4-5% rise in domestic electricity consumption, contributing to increased energy instability. This conclusion was drawn from the analysis of individual smart cadence data collected from Arizona and Illinois [9]. Sánchez-López et al. conducted an investigation into the evolution of energy demands by analyzing hourly data across domestic, marketable, and artificial demand during the initial surge of COVID-19 [10]. Examining the changes in household hourly electricity demand following the epidemic, especially due to the increase in remote work, provides essential information for electricity system operators concerning operational efficiency and management strategies. Additionally, by analyzing the changes in the spatial and temporal patterns of energy consumption, policymakers can develop more informed strategies to improve the integration of renewable energy sources into the power grid. The analysis of high-frequency electricity data facilitates the comprehension of electricity consumption patterns among distinct consumer groups, especially households that have integrated new technologies such as photovoltaics (PV), batteries, and electric vehicles (EV). Qiu et al. [11] employed a difference-in-differences methodology to examine high-frequency smart cadence data from 1,600 electric vehicle (EV) households, demonstrating that individuals enhanced EV charging during off-peak hours characterized by lower prices.

Al Khafaf et al. [12] performed a comparative analysis of electricity consumption among consumers employing photovoltaic (PV) systems equipped with energy storage systems (ESS) versus those lacking ESS. The analysis utilized data collected from more than 5,000 energy consumers, employing 30-minute interval recordings from smart meters. During periods of high temperatures, the deployment of batteries aids in decreasing peak power usage in the autumn season. Qiu et al. (2022b) analyzed hourly electricity data from Arizona and identified a notable variation in consumption patterns among photovoltaic (PV) consumers after the implementation of battery storage systems. Liang et al. (2022a) presented empirical evidence from Arizona concerning the relinquishment of heat pumps, demonstrating that these systems do not inherently lead to energy savings [14]. The integration of electric vehicle charging profiles with residential electricity data enables the assessment of electric vehicles' effects on electricity distribution networks [15]. The identified patterns enable residents to analyze the beneficial impacts of new technology implementations and explore the extent to which the adoption of these technologies affects the electric grid's capacity. Forecast analysis is contingent upon the data employed in the training process, with high-frequency smart cadence data significantly improving the precision of the prediction model. High-resolution predictive models that employ robust data-driven algorithms necessitate validation using high-frequency data. The recent rise in the adoption of smart measures has generated opportunities for improving household cargo forecasting. Precise forecasting of electricity cargo provides essential theoretical support for the smart grid, encompassing demand response, energy operations, and structural planning and investment. Sousa and Bernardo [16] performed a comparative analysis of the effectiveness of multivariate adaptive regression splines, arbitrary timbers, and artificial neural networks in predicting the cargo for the subsequent day, utilizing half-hourly readings from 5,567 homes.

3. PROPOSED SYSTEM

The proposed system is designed to improve exertion monitoring within smart homes by utilizing IoT-enabled detector networks in conjunction with machine learning (ML) models for real-time exertion classification and anomaly detection. The system utilizes vibrant environmental and wearable sensors to gather data on human conditions, facilitating intelligent automation, anomaly detection, and enhanced decision-making in smart home environments and also aims to improve exertion monitoring within smart homes through the implementation of a Decision Tree Classifier (DTC), and another model with the AdaBoost (Adaptive Boosting) Algorithm. The proposed DTC model as a foundational learner provides enhanced precision, resilience, and stability in real-time activity recognition and anomaly detection.

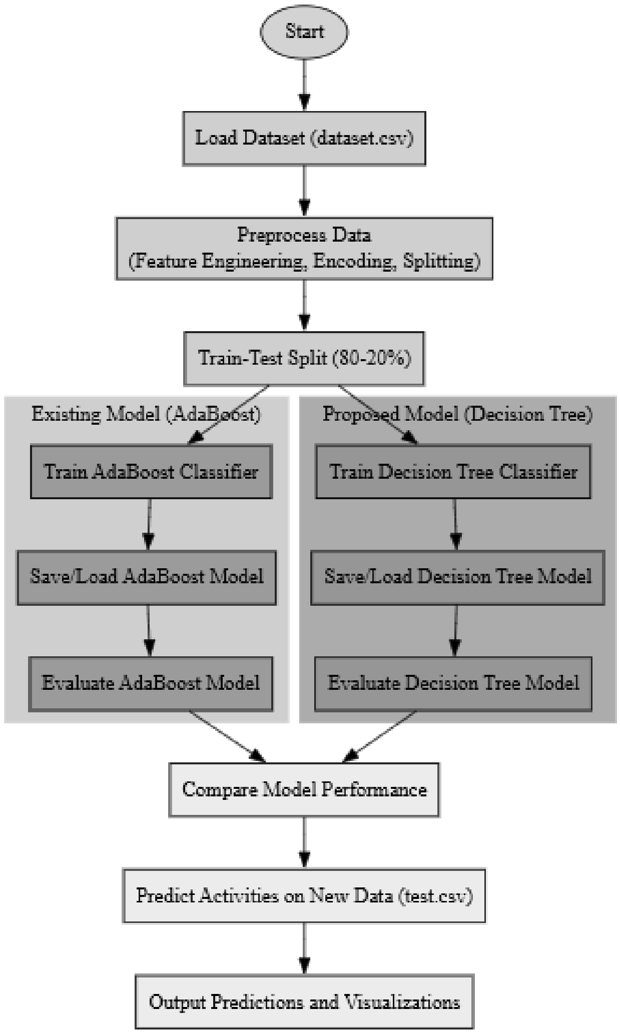

Fig. 2: Proposed system architecture.

Data Collection: IoT Sensors in Smart Homes IoT-enabled detectors capture real-time, timestamped exertion data from various smart home devices. The data encompasses parameters such as movement, temperature, light conditions, door access logs, and appliance operation, among others.

Data Preprocessing and Feature Engineering:

- Transform timestamp data into relevant features such as Month, Hour, Minute, Second, and others.
- Utilize LabelEncoder to transform categorical variables (exertion markers).
- Address the absence of values and eliminate discrepancies. Divide the dataset into training and testing sets, allocating 80% for training and 20% for testing.

DTC model for Activity Recognition: Constructs a hierarchical structure to categorize conditioning. Divides data at irregularities based on the importance of specific points. The workshop effectively utilizes structured data; however, it may be prone to overfitting issues.

- **Simplicity & Interpretability**: Decision trees are easy to interpret, visualize, and deploy.
- **Faster Training Time**: Unlike AdaBoost, which combines multiple weak classifiers, a single decision tree is computationally less expensive.
- **No Need for Boosting**: A well-optimized decision tree can achieve comparable performance without requiring iterative weight adjustments.
- **Handling of Feature Importance**: Decision trees inherently provide insights into the most important features influencing the classification.

4. RESULTS AND DISCUSSION

Figure 3 represents a structured dataset containing smart home sensor data used for activity recognition. The table includes multiple binary features (0s and 1s) corresponding to the state of various home elements such as doors, carpets, lamps, beds, and bathroom-related activities. Each row represents a recorded instance at a specific timestamp, with the "Activity" column indicating the detected human activity (e.g., "sleep" or "anomaly"). The dataset suggests that sensor activations are monitored over time to classify normal activities and detect anomalies, likely for an intelligent home automation or security system.

mainDoor	...	bedroomDoor	bedroomCarp	bedTableLamp	bed	bathroomLight	bathroomDoorLock	bathroomDoor	bathroomCarp	Activity	timestamp
0	...	0	1	0	1	0	0	0	0	sleep	2021-03-01 07_55_16
0	...	0	1	0	1	0	0	0	0	sleep	2021-03-01 07_55_17
0	...	0	1	0	1	0	0	0	0	sleep	2021-03-01 07_55_18
0	...	0	1	0	0	0	0	0	0	sleep	2021-03-01 07_55_19
0	...	0	1	0	0	0	0	0	0	sleep	2021-03-01 07_55_20
...	...	...	...	...	...	...	...	...	...	...	...
0	...	0	0	0	0	0	0	0	0	anomaly	2021-03-14 23_24_23
0	...	0	0	0	0	0	0	0	0	anomaly	2021-03-14 23_24_24
0	...	0	0	0	0	0	0	0	0	anomaly	2021-03-14 23_24_25
0	...	0	0	0	0	0	0	0	0	anomaly	2021-03-14 23_24_26
0	...	0	0	0	0	0	0	0	0	anomaly	2021-03-14 23_24_27

Fig. 3: Uploading Dataset

Figure 4 is a bar chart titled "Count Plot," representing the distribution of different activity categories based on their recorded counts. The x-axis lists categories such as "sleep," "eat," "leisure," "other," "personal," "anomaly," and "work," while the y-axis represents the count of occurrences for each category. The "anomaly" category has the highest count (408,683), followed by "sleep" (273,386), "leisure" (155,414), and "other" (155,352). The "eat," "personal," and "work" categories have significantly lower counts, with "work" being the least frequent (10,835). The plot visually highlights the distribution of activities, suggesting a focus on anomaly detection in behavioral patterns.

Figure 5 presents a confusion matrix corresponding to an AdaBoost Classifier model employed in a multi-class bracket problem. This model categorizes different exertion orders, which encompass sleep, work, rest, personal activities, eating, other tasks, and anomalies. The confusion matrix is represented as a heatmap, with darker shades indicating higher values, thereby emphasizing regions of concentrated predictions. The slant values indicate precise groupings, while the out-slant values represent misclassifications. The classifier successfully identified 28,659 instances of rest, 15,888 instances of other, and 58,043 instances of sleep, indicating satisfactory performance across these categories. There are notable instances of misclassification, including 45,517 cases categorized as anomalies and 14,299 cases incorrectly identified as "sleep" under the "other" category. The work order included a total of 2,192 accurate groups, suggesting a possible difficulty in distinguishing it from other orders. The presence of misclassifications highlights particular aspects where enhancements to the AdaBoost model may be advantageous, potentially through the integration of a more robust classifier, such as a Decision Tree, as recommended in this design.

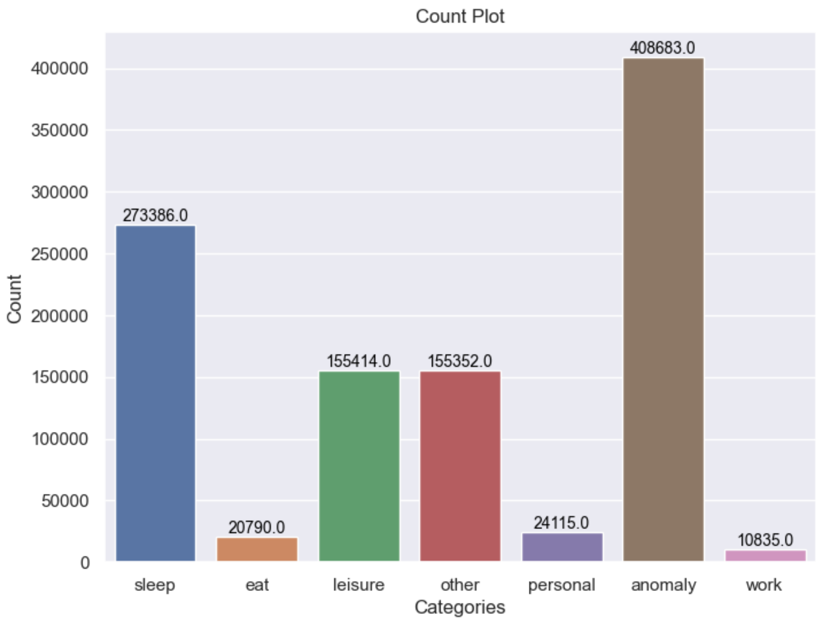

Fig. 4: Count plot.

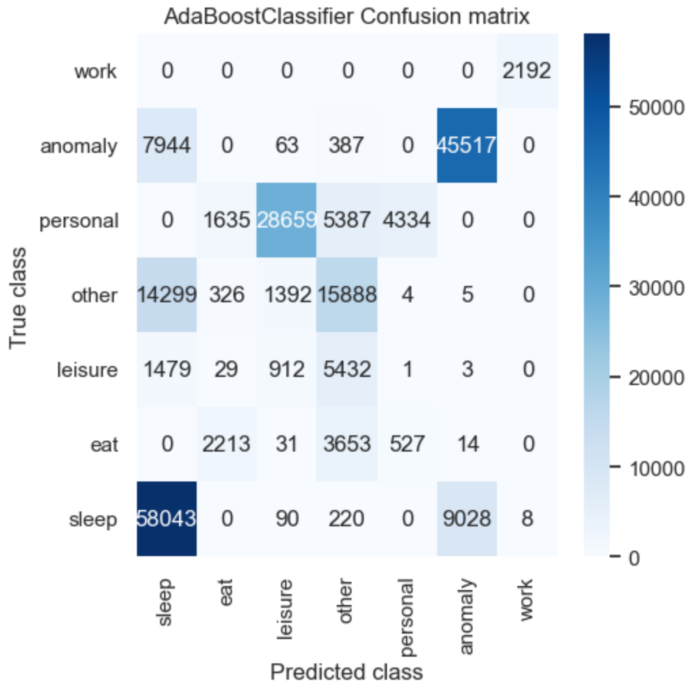

Fig. 5: Confusion matrix obtained using AdaBoost classifier.

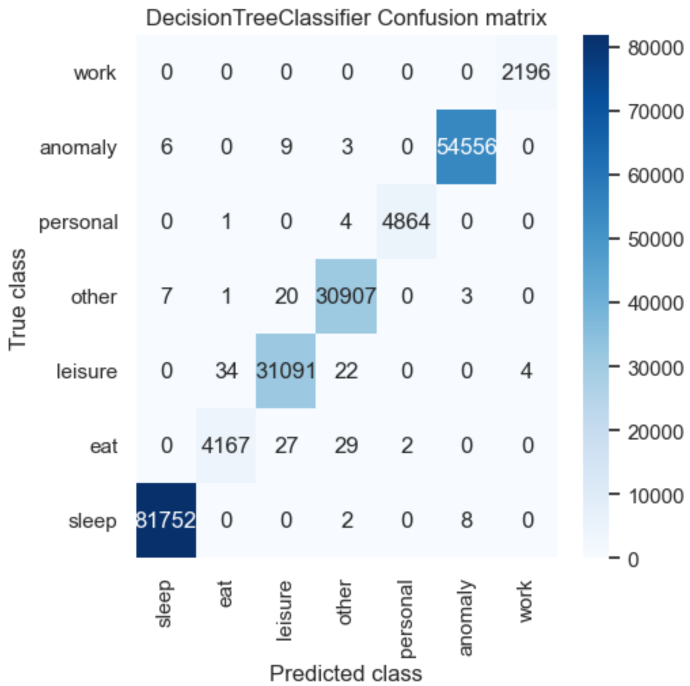

Fig. 6: Confusion matrix obtained using DTC model.

officeCarp	office	mainDoorLock	mainDoor	...	bathroomDoorLock	bathroomDoor	bathroomCarp	Month	Hour	Minute	Day	Year	Second	predict
0	0	1	0	...	0	0	0	3	7	55	1	2021	2021	Sleep
0	0	1	0	...	0	0	0	3	7	55	1	2021	2021	Sleep
0	0	1	0	...	0	0	0	3	7	55	1	2021	2021	Sleep
0	0	1	0	...	0	0	0	3	7	55	1	2021	2021	Sleep
0	0	1	0	...	0	0	0	3	7	55	1	2021	2021	Sleep
0	0	1	0	...	0	1	1	3	13	24	2	2021	2021	Other
0	0	1	0	...	0	1	1	3	13	24	2	2021	2021	Other
0	0	1	0	...	0	1	1	3	13	24	2	2021	2021	Other
0	0	1	0	...	0	1	1	3	13	24	2	2021	2021	Other
0	0	1	0	...	0	1	1	3	13	24	2	2021	2021	Other
0	0	1	0	...	0	1	1	3	13	24	2	2021	2021	Other
0	0	1	0	...	1	0	1	3	12	5	3	2021	2021	Anomaly
0	0	1	0	...	1	0	1	3	12	5	3	2021	2021	Anomaly
0	0	1	0	...	1	0	1	3	12	5	3	2021	2021	Anomaly
0	0	1	0	...	1	0	1	3	12	5	3	2021	2021	Anomaly
0	0	1	0	...	1	0	1	3	12	5	3	2021	2021	Anomaly
0	0	1	0	...	1	0	1	3	12	5	3	2021	2021	Anomaly
0	0	1	0	...	0	0	0	3	19	58	3	2021	2021	Leisure
0	0	1	0	...	0	0	0	3	19	58	3	2021	2021	Leisure
0	0	1	0	...	0	0	0	3	19	58	3	2021	2021	Leisure
0	0	1	0	...	0	0	0	3	19	58	3	2021	2021	Leisure
0	0	1	0	...	0	0	0	3	19	58	3	2021	2021	Leisure
0	0	1	0	...	0	0	0	3	16	6	4	2021	2021	Personal

Fig. 7: Predicted output on test data

Figure 6 illustrates a confusion matrix associated with a DTC model, detailing the classification performance across various activity categories, including sleep, work, leisure, personal, eat, other, and anomaly. The matrix is represented through a heatmap, with darker shades denoting elevated values. The Decision Tree model exhibits enhancements compared to the AdaBoost classifier in accurately categorizing instances such as "other" (30,907 correct classifications) and "leisure" (31,091 correct classifications), which were previously associated with significant misclassifications. In a similar manner, the classification of "sleep" has been accurately achieved 81,752 times, indicating an enhancement relative to the performance of AdaBoost. The category labeled "anomaly" demonstrates improved classification accuracy, with 54,556 correct instances recorded, thereby minimizing instances of misclassification. Nonetheless, certain misclassifications persist, especially concerning the term "eat," which accounts for 4,167 misclassified instances, along with minor inaccuracies in categories such as "personal" and "work." The DTC model demonstrates enhanced accuracy in differentiating various activity categories, thereby highlighting its potential superiority compared to AdaBoost in this project. Finally, Fig. 7 demonstrate the sample predictions on new test data using the proposed DTC model.

5. CONCLUSION

The project involved an analysis of the performance metrics associated with AdaBoost and DTC models in the context of classifying multiple activity categories, which include sleep, work, leisure, personal, eat, other, and anomaly. The AdaBoost classifier demonstrated effectiveness in certain domains; however, it showed considerable misclassification, especially within the "other" and "personal" categories, resulting in ambiguity among various activities. The DTC model exhibited superior classification accuracy, leading to a notable enhancement in the identification of categories such as "leisure," "other," and "sleep." The confusion matrix for the Decision Tree model indicated a significant decrease in misclassification errors, particularly concerning anomalous activities, thereby establishing it as a more appropriate option for this task. Our findings indicate that a classification model based on Decision Trees demonstrates greater effectiveness compared to AdaBoost for this dataset. The enhancements observed in classification accuracy and the reduction in errors demonstrate that Decision

Trees are more effective in identifying the key characteristics of activity categories. Future work may concentrate on the optimization of hyperparameters associated with the Decision Tree or the integration of ensemble methods, such as Random Forest, to improve classification performance further.

REFERENCES

[1] 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

[2] 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

[3] 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

[4] 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

[5] 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

[6] 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

[7] 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

[8] 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

[9] 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

[10] 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

[11] Liang, J., Qiu, Y., and Xing, B. (2022b). Impacts of the co-adoption of electric vehicles and solar panel systems: empirical evidence of changes in electricity demand and consumer behaviors from household smart meter data. Energy Econ. 112:106170. doi: 10.1016/j.eneco.2022.106170

[12] Sousa, J. C., and Bernardo, H. (2022). Benchmarking of load forecasting methods using residential smart meter data. Appl. Sci. 12:9844. doi: 10.3390/app12199844