ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar233

Anomaly Detection in Kitchen Appliance Usage using Machine Learning Approach

Javid Gahramanov[1]; Naeem Naseer[2]; Ayishagul Gahramanova[3]

[1]Software Architect, JVD Smart Systems, Baku, Azerbaijan
[2]Lecturer, Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan
[3]Mathematician Baku, Azerbaijan

Publication Date: 2025/03/20

Abstract: Anomaly detection in domestic energy consumption is crucial for improving energy efficiency, identifying faulty appliances, and identifying unusual patterns of consumption. In this paper, the Isolation Forest (IF) algorithm is employed to detect anomalies in domestic appliance power consumption. The data consists of high-resolution (1Hz) domestic appliance power measurements with three chosen features: DISHWASH, HEATHOME, and AIRCOND, corresponding to dishwasher consumption, home heating, and air conditioning, respectively. These features were selected because they are major contributors to overall energy usage and have the ability to point towards unusual trends. Preprocessing involved normalizing the selected features and using Isolation Forest for anomaly detection. IF separates anomalies by recursive partitioning, effectively separating outliers from regular data. KDE analysis identified that all three features have a bimodal distribution, which indicates different consumption patterns. Pair plot and 3D visualization outcomes verify that IF accurately detected two separate outliers. Further, the histogram of IF anomaly scores verify these results, wherein a 5% contamination level is utilized to distinguish normal and anomalous points. The findings show that unsupervised machine learning models such as Isolation Forest are able to identify effectively anomalies in household energy consumption, providing information on potential energy losses and system errors. The study aims to enhance smart energy monitoring systems through predictive maintenance and energy optimization applications.

Keywords: *Anomaly Detection, Isolation Forest, Energy Consumption, Machine Learning, Smart Home Monitoring.*

How to Cite: Javid Gahramanov; Naeem Naseer; Ayishagul Gahramanova (2025). Anomaly Detection in Kitchen Appliance Usage using Machine Learning Approach. *International Journal of Innovative Science and Research Technology*, 10(3), 425-430. https://doi.org/10.38124/ijisrt/25mar233

I. INTRODUCTION

The past few years, technologies in smart home sensors have developed significantly and have been able to capture and exchange information more efficiently. Applications of artificial intelligence and machine learning (ML) have made the outcomes more accurate and solid. This has enabled a lot of applications to perform intricate calculations with ease. One of them is human activity recognition (HAR), particularly recognizing abnormal user behavior in the home [1].

Anomaly detection is a method that defines any irregular occurrence in a dataset [2]. In the context of a smart home, this method can be of great interest since one's daily routine usually represents their physical and mental health. Through the monitoring of these routines, the system may identify deviations from standard activities, which can suggest changes in health or cognitive function, such as the onset of dementia [3]. This technology is most useful for people who do not want to stay at home but go to healthcare facilities, such as the elderly [4].

A typical anomaly detection system comprises two major steps. The first step is to collect data from the real-world environment with sensors, either vision-based or sensor-based, depending on the domain [5]. The second phase is the behavior analysis phase where anomalous behaviors are identified and proper ML methodologies are utilized for distinguishing normal behavior from anomalous behavior. Anomalous behavior commonly arises when an activity is undertaken at an unknown time [6] or in an unrecognizable frequency and pattern.

There are three machine learning (ML) strategies to detect anomalies, namely, supervised, unsupervised, and semi-supervised learning [7]. The selection of the technique is based on the type of dataset. Supervised learning, which uses labeled data, is the methodology used in this research. ML tasks typically come in two flavors: classification and regression [8]. Classification is concerned with discrete outputs, whereas regression estimates continuous values [9]. Binary classification involves the model producing one of two values, e.g., true/false or 0/1[10]. Because this research classifies behavior into either "normal" or "abnormal," it is applicable to a binary classification problem. In this case,

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar233

"abnormal" is used to denote aberrations in smart home appliance and kitchen activity use.

Various studies have investigated the use of supervised ML methods for anomaly detection in smart home settings, resulting in notable contributions to the field [11]. This study extends prior work by tackling an important challenge: dependence on various activities for detecting anomalies. The objective of this research is to identify unusual behavior in a smart home from data collected from smart kitchen appliances.

II. RELATED WORK

Technological improvements in artificial intelligence and machine learning have greatly improved anomaly detection in smart home settings. Researchers have investigated numerous machine learning methods to examine sensor data and detect abnormal behavior patterns. The research aims to enhance the accuracy and efficiency of abnormal activity detection, especially in health monitoring and assisted living applications. This part is a summary of the previous research on machine learning-based anomaly detection, and its major methodologies, challenges, and contributions.

Jakkula and cook [12] utilized a One-Class Support Vector Machine (OC-SVM) for anomaly detection in home automation, leveraging the Weka tool [13]. Sensor readings were obtained from major areas, the dining room, kitchen, and living room, to observe behavioral patterns. The OC-SVM model performed well, with 0.5 as the type I error rate and 1 as perfect score for precision, recall, and F-measure. But the model was trained on normal behavior instances only, i.e., type II error was not indicated, which may affect its reliability in actual usage.

Likewise, Palaniappan, Bhargavi, and Vaidehi [14] proposed a method to identify abnormal activities by iteratively removing potential normal activities based on a state transition table and a multi-class SVM. In Java implementation, their SVM-based system successfully minimized computational time using a kernel function for training. The research incorporated nine various activities, and in order to handle the vast amount of raw data, the sliding window technique was utilized to divide the dataset and regulate the data flow rate. The preprocessing procedure helped to obtain a total accuracy of 94.4%, which testifies to the successfulness of the presented approach in detecting deviations in everyday activities in a smart home environment.

Ardebili, Eken, and Küçük [15], and Han, Yang, and Huang [16] investigated several machine learning algorithms, such as Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), and k-Nearest Neighbors (KNN), to compare their performance for anomaly detection. The comparison among these models offered insights into system stability and real-world applicability.

Ardebili et al. [16] used sensor-based data in the first study for anomaly detection and achieved accuracy rates of 73% using SVM, 87% using RF, 79% using DT, and 68% using KNN. The results showed that RF performed better compared to other models in classification accuracy and thus holds potential for real-world implementation. Han et al. [15] applied a vision-based sensor system for fall detection, comparing the same ML methods. Their findings exhibited greater precision, at 96% for SVM, 93% for RF, 88% for DT, and 92% for KNN. The performance of vision-based systems being better indicates their ability to detect crucial events, like falls, in smart homes. In addition, Elhadad and Tan [17] employed RF and KNN-based anomaly detection system with special emphasis on two features: time sequence and activity sequence. They obtained an accuracy of 85% in their model, thus proving the relevance of sequential data in detecting behavioral anomalies in smart home applications.

III. METHODOLOGY

Herein, applied an anomaly detection method via machine learning models to detect suspicious patterns in home energy consumption. Our dataset has three important features: DISHWASH, HEATHOME, and AIRCOND, which are the most pertinent attributes to select concerning energy usage patterns. Our method is structured around a pipeline format with the steps of data preprocessing, standardizing features, Isolation Forest and Local Outlier Factor (LOF)-based anomaly detection, and visualizing results. The performance of the models was tested with anomaly distribution analysis, model agreement, and visualization of anomaly scores.

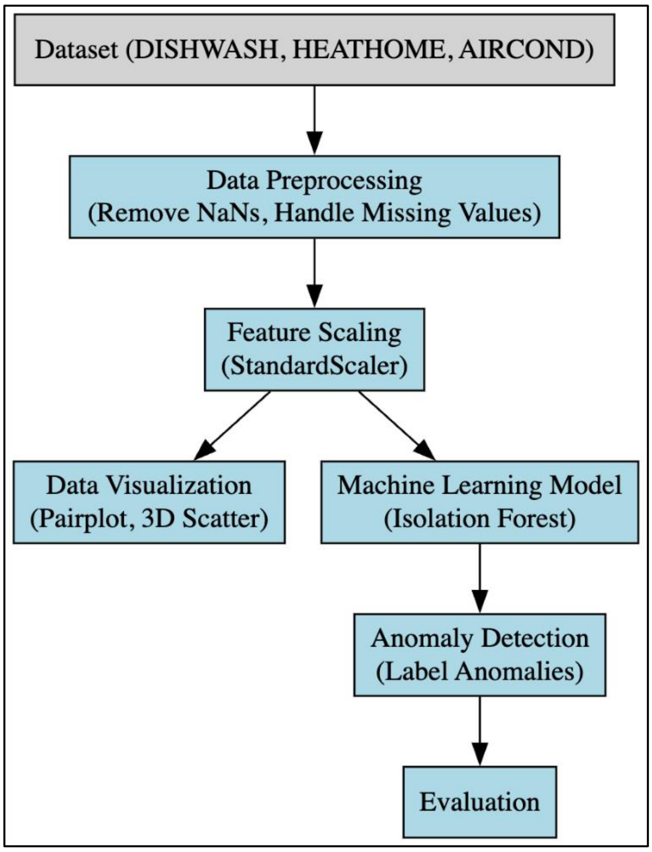

Fig 1 Poposed Methodlogy

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar233

➢ *Dataset Description:*

The data contains residential appliance power consumption recorded at a timestamp resolution of 1Hz. The timestamp field is the date and time in Excel's serial day format, counting days from January 1, 1900, at 12:00 AM. This can be reformatted to show regular date and time values through Excel's number formatting. The other columns include power consumption data for different domestic appliances. Two columns are used for each appliance: one column for measured values (e.g., DP2_Condenser_Power_Val (Actual Values)) and a second column that shows the unit of measurement (e.g., DP2_Condenser_Power_Units). The majority of power values are in Watts. Three special columns with VAR_Val instead of Power_Val hold reactive power measurements, and their units are listed in VARs instead of Watts. The column names provide the names of the respective appliances. The data comes from home load measurements on May 2, 2016.

➢ *Selected Features:*

Three significant fetures DISHWASH, HEATHOME, and AIRCOND in this research work to study anomalies in domestic energy consumption. DISHWASH indicates dish washer usage patterns, HEATHOME represents heating system activity, and AIRCOND represents air conditioning usage. These features were used because they contribute significantly to overall energy consumption and have the capability to indicate unusual patterns like irregular or over usage, which might be indicative of malfunctioning appliances or aberrant behavior. By using anomaly detection methods, we intend to determine deviations that could be used for maximizing energy efficiency and system faults detection. Below Figure Shows the features distribution.

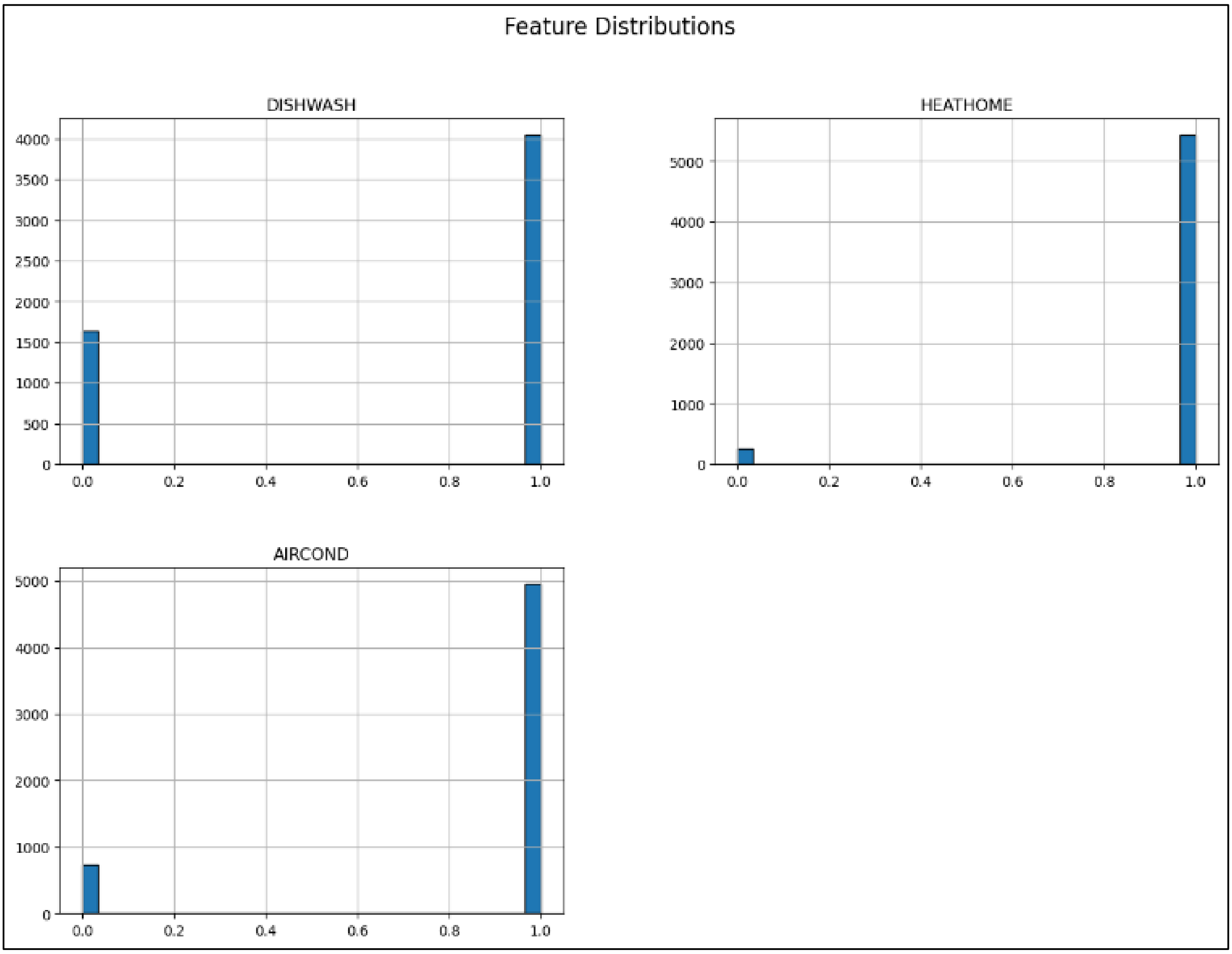

Fig 2 Feature Distribution of Selected Features

The below plot displays Kernel Density Estimation (KDE) plots, showing the distribution of three attributes: DISHWASH, HEATHOME, and AIRCOND. Each of these attributes has a bimodal distribution, which implies two different groups in the data. There is a smaller peak close to 0 for every feature, whereas a broader, higher peak exists at around 1. HEATHOME has the maximum density at the higher peak, followed by AIRCOND and then DISHWASH. The shapes of the distributions are quite comparable, pointing towards possible correlation or common underlying determinants for these features.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar233

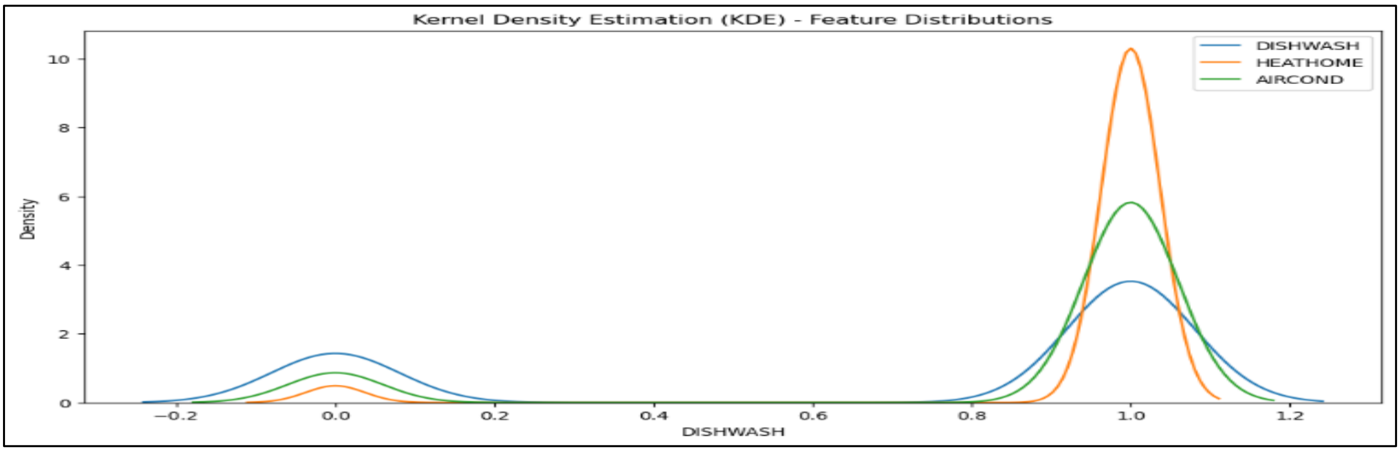

Fig 3 KDE Method

Isolation Forest (IF) is an unsupervised anomaly detection algorithm that separates anomalies rather than profiling typical instances. It operates by randomly choosing one feature and then dividing the data at a random value along that feature. As anomalies are rare and dissimilar from typical data, they are separated in fewer divisions. This renders Isolation Forest efficient and scalable, even for high-dimensional datasets. The algorithm gives anomaly scores depending on the number of splits required to separate a point lower score for anomalies. IF is extensively applied in fraud detection, network security, and industrial monitoring because it is fast and efficient in identifying outliers without needing labeled data.

IV. RESULTS

Diagonal plots indicate distribution of every feature where blue identifies normal data points and red does the anomalies. Remarkably, there are merely two anomalies pinpointed in each and every feature, implying the presence of sparse anomalies. The off-diagonal plots indicate scatter plots of each pair of features further highlighting isolation of these two anomalies from other data points within the cluster. The Anomaly_LOF column, probably a Local Outlier Factor value, clearly distinguishes between the outlier-free points with low values close to zero and the two outliers with higher values. Overall, the plot demonstrates how the Isolation Forest algorithm successfully singles out these two data points as individual outliers among the rest of the dataset, exhibiting the algorithm's capability to identify anomalies in a high-dimensional feature space.

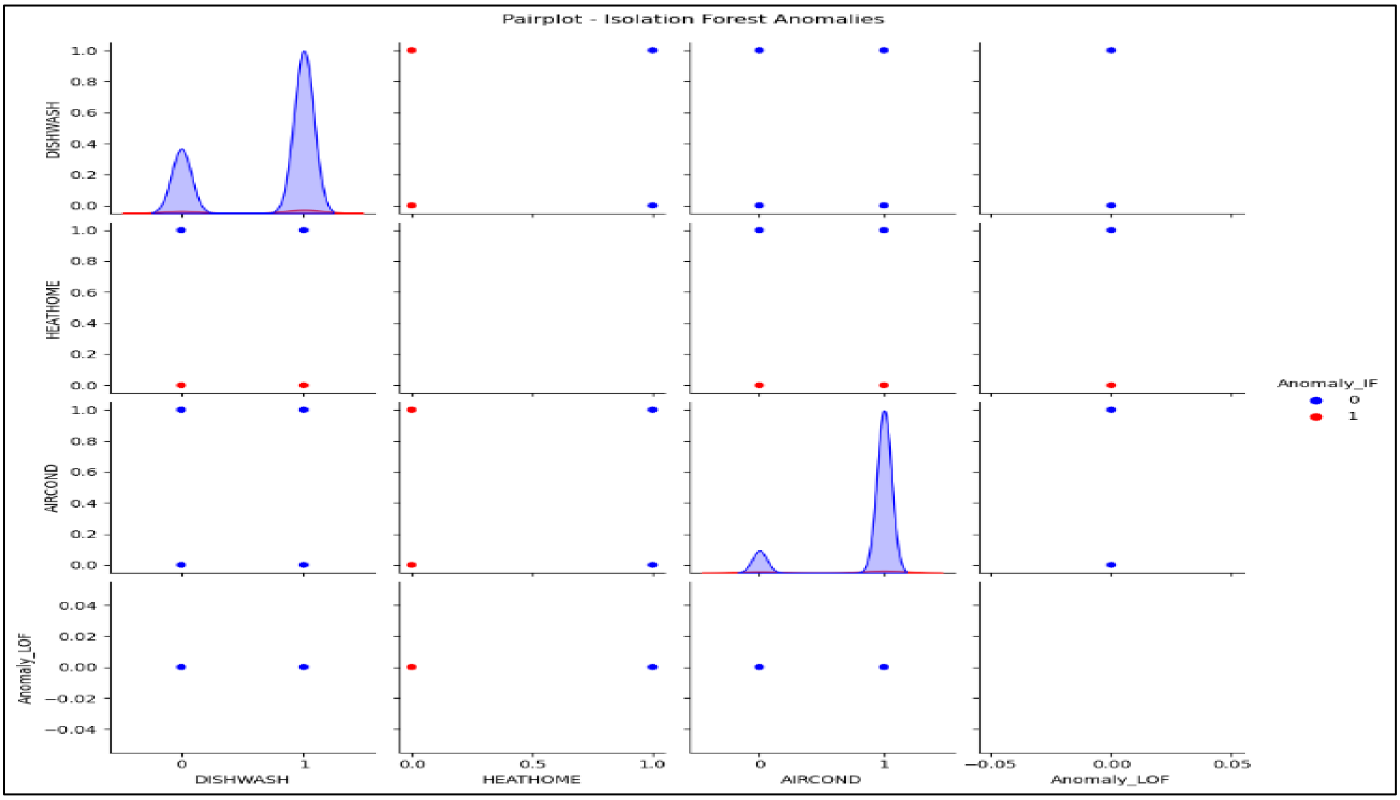

Fig 4 Pair plot for anmoly detection using isolation forest Isolation Forest detected 258 anomalies out of 5686 samples (4.54%)

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar233

The below histogram illustrates the distribution of anomaly scores obtained from an Isolation Forest algorithm, a technique for outlier detection. The x-axis is the anomaly score, with lower scores meaning higher probability of being an anomaly. The y-axis indicates the frequency of data points within each range of scores. The graph is overlaid with a Kernel Density Estimation (KDE) curve, a smoothed view of the distribution of scores. A dashed red vertical line indicates the 5% level, which is the boundary below which data points are defined as anomalies. The histogram shows a bimodal distribution with a dense cluster of data points with scores close to 0.15, which represents normal data, and a sparse cluster with scores lower than the 5% level, which represents possible anomalies. The existence of data points with scores well below zero indicates that the algorithm has picked up on obvious outliers. The 5% threshold does a good job of distinguishing between the bulk of the distribution and the lower-scoring outliers, showing the algorithm's capability to detect deviations from the typical data patterns.

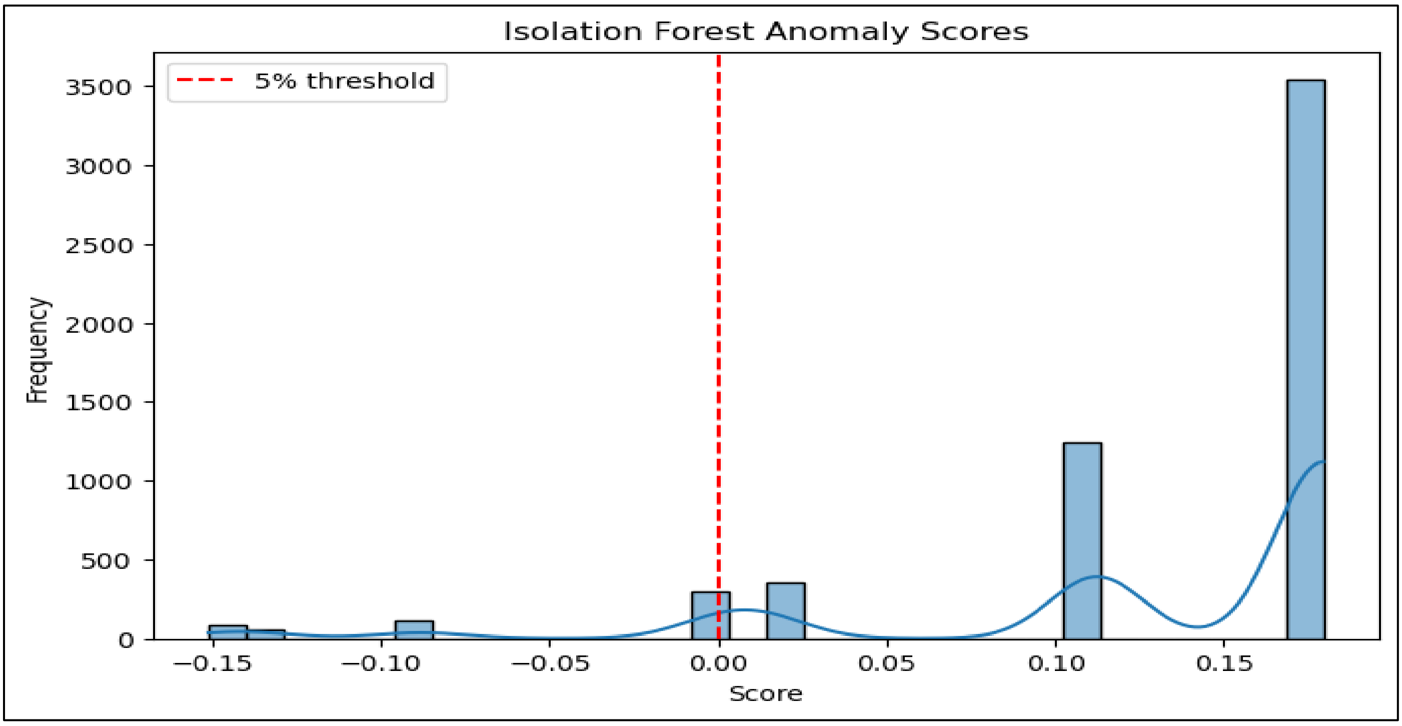

Fig 5 Vissualization of Anomly Detection Score

V. CONCLUSION

This research proved the efficacy of the Isolation Forest (IF) algorithm in identifying anomalies in residential energy consumption, particularly on three of the most significant household appliances: dishwasher (DISHWASH), home heating (HEATHOME), and air conditioning (AIRCOND). The findings show that IF effectively detected abnormal energy consumption patterns, which may indicate possible energy losses, malfunctioning appliances, or abnormal user behavior. Kernel Density Estimation (KDE) analysis confirmed that the features follow a bimodal distribution, which indicates different consumption behaviors. Visual examination using pair plots and 3D visualizations also confirmed that the model could identify outliers. The IF anomaly score histogram also confirmed the results, verifying the robustness of the methodology. Through the use of unsupervised machine learning methods such as Isolation Forest, intelligent home monitoring systems can be used to optimize energy efficiency, facilitate predictive maintenance, and aid in automated fault detection. Additional work may consider the integration of other appliance data, improving model performance through hybrid methods, and integrating real-time anomaly detection to provide proactive energy management

ACKNOWLEDGMENT

Javid Gahramanov expresses gratitude to JVD Smart Systems for their support in facilitating this research. Naeem Naseer acknowledges Muhammad Nawaz Sharif University of Agriculture, Multan, Pakistan, for providing an academic environment conducive to this study. Ayishagul Gahramanova extends appreciation to her colleagues in the field of mathematics for their valuable insights. The authors also thank all contributors and institutions who provided data and resources for this research on anomaly detection in kitchen appliance usage.

REFERENCES

[1]. E. Hoque, R. F. Dickerson, S. M. Preum, M. A. Hanson, A. T. Barth, and J. A. Stankovic, “Holmes: A Comprehensive Anomaly Detection System for Daily In-home Activities,” 2015 International Conference on Distributed Computing in Sensor Systems, pp. 40–51, 2015, [Online]. Available: https://api.semanticscholar.org/CorpusID:15978673
[2]. V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” ACM computing surveys (CSUR), vol. 41, no. 3, pp. 1–58, 2009.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25mar233

[3]. K. Mengoudi et al., "Augmenting dementia cognitive assessment with instruction-less eye-tracking tests," IEEE J Biomed Health Inform, vol. 24, no. 11, pp. 3066–3075, 2020.
[4]. A. C. Tran, S. Marsland, J. Dietrich, H. W. Guesgen, and P. Lyons, "Use cases for abnormal behaviour detection in smart homes," in Aging Friendly Technology for Health and Independence: 8th International Conference on Smart Homes and Health Telematics, ICOST 2010, Seoul, Korea, June 22-24, 2010. Proceedings 8, Springer, 2010, pp. 144–151.
[5]. L. M. Dang, K. Min, H. Wang, M. J. Piran, C. H. Lee, and H. Moon, "Sensor-based and vision-based human activity recognition: A comprehensive survey," Pattern Recognit, vol. 108, p. 107561, 2020.
[6]. N. C. Tay, T. Connie, T. S. Ong, A. B. J. Teoh, and P. S. Teh, "A review of abnormal behavior detection in activities of daily living," IEEE Access, vol. 11, pp. 5069–5088, 2023.
[7]. U. Fiore, F. Palmieri, A. Castiglione, and A. De Santis, "Network anomaly detection with the restricted Boltzmann machine," Neurocomputing, vol. 122, pp. 13–23, 2013.
[8]. S. Omar, A. Ngadi, and H. H. Jebur, "Machine learning techniques for anomaly detection: an overview," Int J Comput Appl, vol. 79, no. 2, 2013.
[9]. P. Harrington, Machine learning in action. Simon and Schuster, 2012.
[10]. Y. Wang et al., "A survey on ambient sensor-based abnormal behaviour detection for elderly people in healthcare," Electronics (Basel), vol. 12, no. 7, p. 1539, 2023.
[11]. A. B. Nassif, M. A. Talib, Q. Nasir, and F. M. Dakalbab, "Machine learning for anomaly detection: A systematic review," Ieee Access, vol. 9, pp. 78658–78700, 2021.
[12]. V. Jakkula, "Predictive data mining to learn health vitals of a resident in a smart home," in Seventh IEEE International Conference on Data Mining Workshops (ICDMW 2007), IEEE, 2007, pp. 163–168.
[13]. M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," ACM SIGKDD explorations newsletter, vol. 11, no. 1, pp. 10–18, 2009.
[14]. A. Palaniappan, R. Bhargavi, and V. Vaidehi, "Abnormal human activity recognition using SVM based approach," in 2012 international conference on recent trends in information technology, IEEE, 2012, pp. 97–102.
[15]. K. Han, Q. Yang, and Z. Huang, "A two-stage fall recognition algorithm based on human posture features," Sensors, vol. 20, no. 23, p. 6966, 2020.
[16]. E. Seyedkazemi Ardebili, S. Eken, and K. Küçük, "Activity recognition for ambient sensing data and rule based anomaly detection," The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 44, pp. 379–382, 2020.
[17]. R. M. Elhadad and Y.-F. Tan, "Anomaly Detection for Human Home Activities Using Pattern Based Sequence Classification.," Journal of ICT Research & Applications, vol. 17, no. 1, 2023.