



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** X **Month of publication:** October 2026

DOI: <https://doi.org/10.22214/ijraset.2026.85122>

www.ijraset.com

Call: ☎ 08813907089

E-mail ID: ijraset@gmail.com

ML-Based Smart Appliance Energy Monitoring System

Divya Dongre¹, Astha Bavistale², Atharva Wasnik³, Rohan Hedau⁴, Sanchit Nardelwar⁵, Krishnakumar Bhabhda⁶, Ms. Priya Meshram⁷

Department of Computer Science and Engineering, G.H Rasoni University, Amravati, India

Abstract: Modern homes only get a combined electricity consumption value. In other words, they have no information about which of their appliances have used the electricity a house has consumed. This paper introduces the design of an ‘ML-Based Smart Appliance Energy Monitoring’ system for recognizing the electricity consumption of a smart home. The Non-Intrusive Load Monitoring system uses an ESP32 development board and a PZEM-004T V3.0 energy meter and current transformer (CT) to measure the voltage, current, power, frequency, and power-factor data of a particular circuit of the home’s electrical circuit. These are then transmitted over Wi-Fi/HTTP to a Flask-based inference server where a 50-feature input (constructed in identical fashion between training and inference from a 13,439-row labeled dataset) is classified by a joint ensemble (HistGradientBoosting (45%), ExtraTrees (35%), and Random Forest (20%)) of regressors into one of eight appliance-combination classes representing three target devices (a phone charger, a laptop charger, and a hot glue gun). An additional individual-device model enables ON-time tracking, per-device energy estimation, and electricity cost calculation for the devices. Classification results are then locally visualized on a 16 × 2 I2C LCD and remotely on the Blynk IoT platform while storing relevant device statistics in the ESP32 Preferences. In contrast to the recent deep-learning-based approaches, which typically require heavy computation and a large amount of annotated data to train on, we propose an explainable ensemble pipeline for low-end single-household appliances.

Keywords: Non-Intrusive Load Monitoring (NILM); Energy Disaggregation; Ensemble Machine Learning; Random Forest;

I. INTRODUCTION

Residential electricity use is reported to end users as an aggregate value on a utility meter. While adequate for billing purposes, this offers a household only a global view of the consumption and no information on which appliance is responsible for a given portion. The appliance-level insight is also critical for the purpose of demand-side energy management, which is becoming increasingly important as electricity grids incorporate a higher share of variable renewable generation and require more granular load information to match supply and demand.

Two approaches exist to obtain the appliance-level consumption data. Intrusive Load Monitoring (ILM) deploys a dedicated sensor on every individual appliance, which provides accurate readings but is expensive to install and scale with the number of appliances. Non-Intrusive Load Monitoring (NILM), first introduced by Hart [1], on the contrary, infers the operating state and power draw of the individual appliances from a single measurement point at the main supply line.

The main contribution of this paper is to present a practical NILM pipeline that is realistic to implement, train and evaluate with limited computational and data resources. In particular, this work:

- (i) proposes a single-point hardware sensing design that makes use of a commercial metering module in place of hand-crafted analog sensing circuits;
- (ii) defines a 50-feature extraction pipeline that is kept identical for both training and live inference;
- (iii) employs a weighted joint ensemble of classical, explainable classifiers that distinguish eight appliance-combination classes, supported by a separate individual-device model
- (iv) presents two local and remote interfaces for appliance-level visualization.

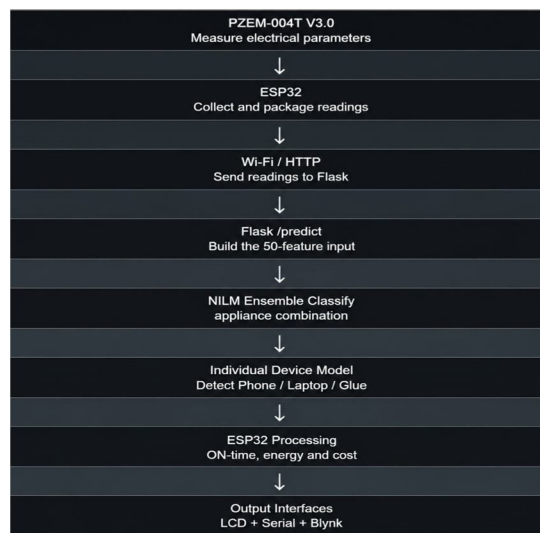


Fig.1 Overall system flowchart of the proposed ML-based smart appliance energy monitoring system

II. RELATED WORK

The problem of Non-Intrusive Load Monitoring (NILM) was initially introduced by Hart who proposed to infer a state of individual appliances from the measured overall electricity consumption instead of installing additional sensors on each appliance [1]. Later, NILM studies have investigated a variety of probabilistic, signal-based processing and machine learning methods seeking to decompose and extract a consumption of each individual appliance in the aggregate measurements.

Recently, the application of deep-learning based approaches is gaining more and more attention, due to their promising results in numerous domains. Kong et al. [2] proposed a hybrid CNN-BiLSTM-Transformer architecture, which is capable to combine the convolutional feature extractor, BiLSTM layers, capturing long-term temporal patterns, and Transformer blocks, which are able to learn the interdependencies between all elements of a sequence. The hybrid model was tested on the UK-DALE and REDD datasets, improving upon the results of the REDD by 24.5% (F1-score) and the results of the UK-DALE by 22.8% (F1-score) compared to the considered baseline models. Even though such a hybrid architecture has a great modeling power, the increased complexity may lead to increased demands for computations and resources, which may be more suitable for high-end applications and not for low-end devices.

Besides standard recurrent architectures for sequence modeling, attention-based mechanisms could be also employed to learn fine-grained representations of the appliance load profiles. Pan et al. [3] proposed an Adaptive Multi-Scale Attention Integration Module (AMSAIM), which takes advantage of the efficient multi-scale attention and adaptive receptive fields to learn appliance-specific characteristics. The method was evaluated on the UK-DALE dataset, achieving average F1-scores of approximately 0.94 for the seen cases, as well as competitive results for unseen cases. electronics-15-00517.pdf However, it was based on a relatively complex deep-learning architecture, and the performance was mostly validated on the UK-DALE.

A systematic review and meta-analysis of the machine-learning-based methods for NILM appeared in [4], where the performance of 30 included studies was quantitatively evaluated in terms of F1-score, published between 2017 and 2024. The study found significant heterogeneity both among the included studies and between the different factors that were considered in relation to the outcome of the experiments. This indicates the importance of the interpretation of the results based on the characteristics of the dataset and the used procedure, instead of emphasizing the obtained accuracy level.

Another promising research direction is represented by methods utilizing transfer learning, which enables to obtain good performance with a relatively small amount of labeled data. In [5], Garcia-Marrero et al. proposed RefQuery, a multi-appliance NILM approach, which is based on a single pretrained model, shared across all appliances, while appliance-specific encodings are represented by lightweight embeddings learned during the target domain adaptation. Thus, RefQuery is able to incorporate information about new target appliances with a relatively small amount of computation. The network was evaluated on three public datasets, demonstrating a reasonable accuracy-efficiency trade-off and enabling the deployment on less powerful edge devices. However, such methods still rely on the existence of some anchor examples of the target domain in order to adapt weights of the pretrained networks.

In a more recent work, Petralia et al. [6] proposed NILM Former, which was specifically designed to address the stationarity assumption violation in real-world smart-meter electricity time series. The method introduces a subsequence stationarization/de-stationarization mechanism, as well as a Transformer-based encoder, which was used to achieve better modeling of the input signals. The effectiveness of the approach was validated on four real-world datasets, demonstrating significantly better performance than the considered state-of-the-art methods. Though NILMFormer is able to effectively model the non-stationary input signals, the utilization of Transformer blocks requires a considerable amount of computations, which limits applicability to devices with low performance.

In general, the above-mentioned studies demonstrate encouraging results in terms of accuracy and scalability, but mostly require significant data to train the models and utilize relatively complex deep-learning architectures, which limits utilization to low-end devices. Instead, in this work, we propose a relatively lightweight and interpretable model based on the histogram gradient-boosting and tree-based algorithms. We explore the application of Hist Gradient Boosting, ExtraTrees and Random Forest regressors for solving an appliance-combination classification task and estimating individual appliance consumption based on the data collected from a single household.

Table I. Comparison Of Related Nilm Approaches

Ref.	Approach / Model	Dataset / Evaluation	Key Contribution	Limitation / Gap
[1]	Traditional NILM (event/signature-based)	Household aggregate measurements (conceptual study)	Introduced the NILM concept of inferring appliance states from aggregate load measurements at a single point.	Limited ability to handle simultaneous appliance operation and complex appliance signatures.
[2]	Hybrid CNN-BiLSTM-Transformer	UK-DALE, REDD	Combines CNN for feature extraction, BiLSTM for long-term dependencies, and Transformer blocks to model complex appliance load patterns. Improves F1-score by 24.5% (REDD) and 22.8% (UK-DALE).	Architecture is more complex and computationally demanding than lightweight machine-learning models.
[3]	Adaptive Multi-Scale Attention Integration Module (AMSAIM)	UK-DALE	Uses multi-scale attention with EMA and SK modules to improve NILM accuracy and real-time performance. Achieves an average F1-score ≈ 0.94 in seen scenarios.	Higher computational complexity and limited validation scenarios (primarily UK-DALE).
[4]	Machine Learning Systematic Review and Meta-Analysis	30 studies from multiple public datasets (2017–2024)	Provides a quantitative comparison of ML models for NILM and analyzes factors affecting performance. Hybrid models generally show higher F1-scores.	Not a new NILM model; performance varies significantly across datasets and experimental settings. High heterogeneity between studies.
[5]	Lightweight and Scalable Transfer Learning Framework (RefQuery)	Three public datasets (REDD, UK-DALE, REFIT)	Proposes a scalable multi-appliance transfer-learning framework with a frozen pretrained model and per-appliance embeddings. Suitable for resource-constrained edge devices.	Performance depends on transfer and adaptation between datasets and appliance domains.
[6]	NILMFormer (Transformer-based)	Four real-world datasets	Introduces a subsequence stationarization/de-stationarization scheme to handle non-stationary smart meter data. Significantly outperforms existing approaches.	Transformer-based architecture remains relatively complex for low-resource implementations.
–	ML-Based Smart Appliance Energy Monitoring System (Proposed)	Self-collected household dataset	Uses 50 engineered electrical features and a weighted ensemble of HistGradientBoosting, ExtraTrees, and Random Forest for appliance-combination detection and individual appliance-estimation, integrated with IoT-based data acquisition.	Currently evaluated on a small set of appliance types; can be extended to more appliances and larger datasets in future work.

III. PROPOSED METHODOLOGY

A. Architecture

The described system is a non-intrusive appliance energy monitoring system where the electrical measurements are acquired at only one single point on the house main supply. Unlike intrusive monitoring systems, which demand the deployment of dedicated sensors per appliance, our approach proposes to exploit aggregate electrical measurements for appliance working condition identification.

The system comprises a data acquisition stage, a stage comprising feature extraction and selection, a classification stage for appliance combinations, individual appliance estimator, and a visualization stage. The electrical values are acquired from the house supply and processed in the processing system, wherein the feature vector is extracted and sent to machine learning models. Finally, the results of these models are used for estimating the appliance-level operation and power information.

The described system aims to provide appliance-level information while keeping the hardware for sensing at the same level as the single-household installation can provide. The complete processing flow of the described system is shown in Fig. 1.

B. Data Collection

An ESP32 development board is connected with a PZEM-004T energy monitoring module and current transformer (CT). The latter is clamped on the live conductor, while the whole circuit is connected to the main line of the household and not the individual appliances. The household supply is connected with an AC input source and AC output socket. The CT is clamped on the live conductor to sense the current that passes through the household supply.

The PZEM-004T module offers the desired electrical measurements of the monitoring system, namely, voltage, current, active power, power factor, and frequency. The measured values are transferred via serial UART from module to Esp32. Since the values are read from the main circuit with supplies, the current consumption of the connected devices can also be displayed on the module without having to equip each appliance to be monitored with a sensor.

Esp32 is connected to 16×2 LCD via I2C bus. The latter offers a local real-time display of the desired electrical parameters that include voltage, current, and power. Thus, the user will have direct access to the acquired parameters at the sensing unit. The data is then transferred via Wi-Fi to the backend server for processing. The backend gets the electrical readings and prepares it for the next step of feature extraction and machine-learning process.

The complete circuit-level wiring diagram of the hardware setup is given in Fig. 2. The hardware consists of the ESP32 development board, the PZEM-004T module, the current transformer, the I2C interface 16×2 LCD display, and the AC input/output sockets connections. The figure below shows the circuit schematic of the hardware configuration.

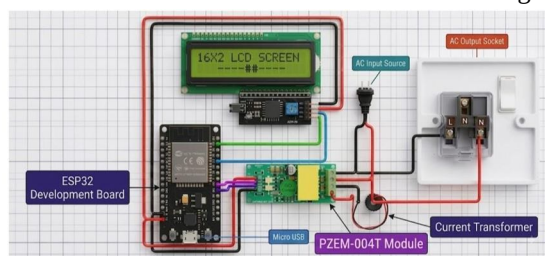


Fig. 2. Circuit-level wiring diagram of the hardware setup.

C. Feature Extraction

The electrical measurements in this section can be converted into a suitable input representation for the machine-learning models. An engineered set of 50 electrical features was found to yield the best performance on the proposed method. The feature-extraction process aims to find a way to convert the raw electrical measurements into features that can be used to distinguish the different operating states of the appliances. In this case, the features are derived from the available electrical measurements and their properties, thus providing more information than would be the case when using only one measurement (e.g., only the power).

Feature engineering plays a particularly crucial role in the case of aggregate-load monitoring, where different combinations of appliances may result in similar power consumption and therefore similar power readings. By using additional features derived from electrical measurements, more information can be provided to help classify the different combinations of simultaneously operating appliances.

The same preprocessing procedure was applied to the training set as will be applied to the incoming input data. This way, the machine-learning models always receive the same type of input representation irrespective of the data set they are trained on or the input data that they process in practice.

D. Appliance-Combination Classification

The extracted feature vector is provided to the appliance-combination classification stage. The proposed system uses a weighted ensemble of three classical machine-learning classifiers, namely, HistGradientBoosting, ExtraTrees, and Random Forest. These are combined to obtain the final classification decision, where the use of multiple classifiers allows the system to combine them rather than rely on a single classifier. The proposed ensemble assigns the following weights to the three classifiers:

- HistGradientBoosting: 45%,
- ExtraTrees: 35%,
- Random Forest: 20%.

The combined weighted predictions are used to identify the appliance-combination class corresponding to the given electrical measurements. The suggested system considers eight appliance-combination classes found in the collected data of the household.

The use of classical ensemble models provides a relatively lightweight machine-learning approach compared with computationally intensive deep-learning architectures. This makes the proposed approach suitable for experimentation and deployment in a resource-constrained single-household environment.

E. Individual Appliance Estimation

The appliance-combination classification stage described above identifies the combination of home appliances that corresponds to the observed aggregate electrical consumption measurements. However, the classification result contains only aggregate information about the state of appliances. In order to infer the state of individual appliances, an additional step of individual-device estimation is needed.

The approach we propose uses a model of an individual device to infer the operating state of individual appliances from the resulting electricity consumption measurements. The appliance individual estimation stage is complementary to the appliance-combination classifier because it allows for finer grained appliance monitoring including operating time tracking and cost estimation. Thus, by combining an appliance-combination classifier and an individual appliance estimator the ultimate goal of obtaining individual appliance information with only whole-house electrical consumption measurements can be achieved.

F. Energy and Cost Estimation

The estimated appliance states can be utilized to obtain the related operating time and energy consumption of appliances. For appliances determined as operating state, the operating time can be recorded by the system.

The energy consumed by an appliance can be expressed as the product of its power and the time that this appliance is working:

$$E = P \times t$$

where E is the energy, P is the power and t is the run time.

The electricity consumption in kWh can be formulated for electricity-consumption analysis. Then the electricity cost of appliances can be estimated by electricity cost based on the electricity tariff :

$$Cost = E_{kWh} \times Tariff$$

Accordingly, it enables the system to provide information about the appliance-level electricity consumption and its related cost.

G. Real-Time Processing and Output Interfaces

Once the electric measurements are achieved, the ESP32 transmits them to the backend server via Wi-Fi to process. The latter performs the feature-extraction process on the acquired measurements, which gives the feature vector required by the machine-learning models.

The resulting feature vector is sent to the weighted ensemble classifier in order to predict the corresponding combination of appliances. The individual-device model is then adopted to get the operating information of the individual appliances. The results are utilized in the estimation of the energy and cost.

The proposed system ensures that the monitored information is accessible from both local and remote locations. The 16×2 LCD displays the local display of the measured electrical parameters whereas the remote interface is set up for the appliance-level monitoring information. Thus, the system provides a practical approach to monitoring without the need for dedicated sensors on all the appliances.

In general, the proposed methodology couples the single-point electrical sensing with engineered feature extraction, classical ensemble machine learning, individual appliance estimation, and energy-cost monitoring to propose a lightweight approach to smart appliance energy monitoring.

An important consideration in non-intrusive load monitoring is the interpretation of aggregate electrical measurements when multiple appliances operate simultaneously. Since the sensing unit measures the combined electrical behavior of the connected loads, the same or similar power values may correspond to different appliance combinations.

IV. RESULTS AND DISCUSSION

A. Model Configuration and Classification Performance

The proposed system was tested as an eight-class appliance-combination classification problem: All Off, Phone Charger, Laptop Charger, Hot Glue Gun, Phone Charger + Laptop Charger, Phone Charger + Hot Glue Gun, Laptop Charger + Hot Glue Gun, and Phone Charger + Laptop Charger + Hot Glue Gun. In total, there were 50 engineered electrical features that were entered into the machine learning model.

Three classifiers, namely, Hist Gradient Boosting, Extra Trees, and Random Forest were tested individually. The proposed weighted ensemble uses the predicted probabilities of the three classifiers with weights of 45%, 35%, and 20%, respectively. The evaluation was performed on a hold-out test set of 13,439 samples. The performance comparison of the evaluated models is shown in Table II.

Model	Accuracy (%)	Precision (Macro) (%)	Recall (Macro) (%)	F1-Score (Macro) (%)
HistGradientBoosting	99.58	99.43	99.50	99.46
ExtraTrees	99.37	99.11	99.35	99.23
Random Forest	99.39	99.14	99.39	99.26
Weighted Ensemble	99.55	99.41	99.48	99.44
Final Prediction after Arbitration	99.34	99.23	99.36	99.29

Fig.3 Performance comparison of the evaluated machine-learning models and the proposed weighted ensemble

The individual classifiers showed that HistGradientBoosting achieved the highest accuracy of 99.58%, together with a macro precision of 99.43%, a macro recall of 99.50%, and a macro F1-score of 99.46%, while ExtraTrees and Random Forest had accuracies of 99.37% and 99.39%, respectively.

The accuracy of the weighted ensemble suggested was 99.55%, the macro precision, recall and F1-score being 99.41%, 99.48% and 99.44% respectively. The fact that these results indicate that the ensemble had performance close to that of the best individual classifier while combining the probability outputs of the three models.

In the final prediction stage a laptop-specific arbitration rule was included as a post-processing procedure and its accuracy amounted to 99.34%, a figure that is less than the 99.55% accuracy achieved earlier than the arbitration was applied. Since the arbitration stage did not enhance the overall hold-out classification accuracy in this instance.

B. Class-Wise Classification Performance

The results for each class show that most of the appliance combinations were identified with both high precision and high recall.

With respect to the HistGradientBoosting classifier, both the All Off and Phone Charger categories obtained a precision and a recall of 100%. The Phone Charger + Hot Glue Gun category also had a recall of 100%. Of the eight classes, the one with the lowest F1-score was Phone Charger + Laptop Charger + Hot Glue Gun, the score in this case being 97.38%.

The recall values for each class in the last prediction following arbitration are given in Table III

Table III . Class-Wise Recall Of The Final Prediction

Appliance Combination	Recall (%)
All Off	100.00
Phone Charger	100.00
Laptop Charger	99.65
Hot Glue Gun	99.91
Phone Charger + Laptop Charger	98.94
Phone Charger + Hot Glue Gun	100.00
Laptop Charger + Hot Glue Gun	98.42
Phone Charger + Laptop Charger + Hot Glue Gun	97.98

The findings indicate that the recall rate for individual appliance states was greater than that for combinations in which multiple appliances were operating at the same time, with the three-appliance combination giving the lowest recall at 97.98%, and the Laptop Charger together with the Hot Glue Gun combination coming next at 98.42%.

The fact is that simultaneous operation leads to more difficult classification conditions since the combined electrical characteristics of the various appliances may overlap. Yet the recall values stay high for all eight classes, which shows that the engineered electrical representation contains strong discriminatory information for classifying appliance combinations.

C. Confusion Matrix Analysis

The confusion matrices were analyzed using the eight appliance-combination classes, with rows corresponding to the actual classes and columns to the predicted classes.

The HistGradientBoosting classifier resulted in a largely diagonal confusion matrix. The 800 All Off samples and all 1,532 Phone Charger samples were correctly classified. Most of the remaining samples were correctly classified too, with only a few being misclassified, primarily in the classes involving the Laptop Charger and simultaneous operation.

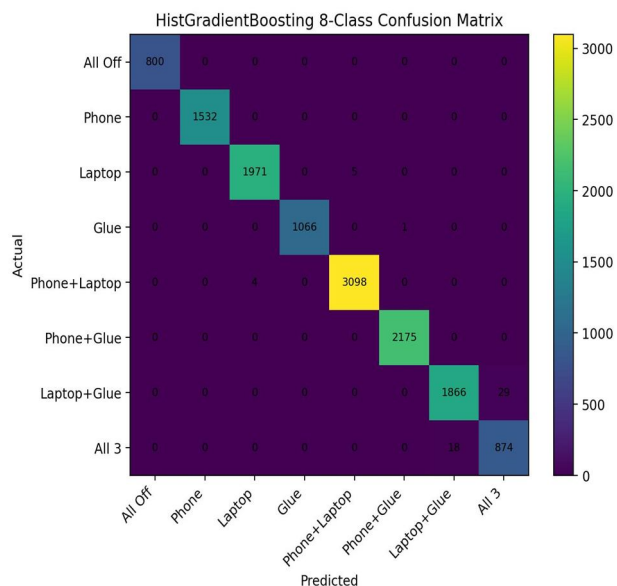


Fig. 4. Confusion matrix of the HistGradientBoosting classifier for the eight appliance-combination classes.

The proposed weighted ensemble also produced a mostly diagonal confusion matrix. Overall, the ensemble correctly classified 13,379 out of the total 13,439 test instances, which gives an accuracy of about 99.55%.

Most often, the misclassified instances came from the similar multi-appliance classes. Indeed, there were 30 instances of the Laptop Charger + Hot Glue Gun class misclassified as the three-appliance combination and 18 instances of the three-appliance class misclassified as the Laptop Charger + Hot Glue Gun combination.

The final weighted-ensemble confusion matrix is presented in Fig. 5.

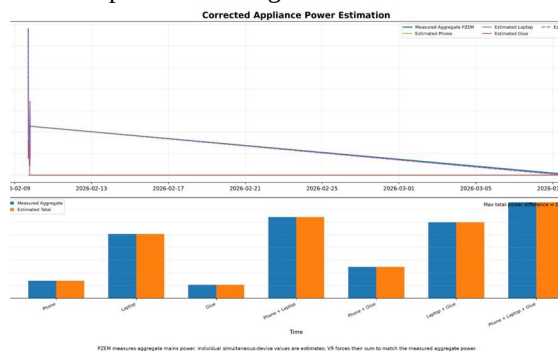


Fig. 5. Confusion matrix of the proposed weighted ensemble classifier for the eight appliance-combination classes.

The concentration of errors between different states of appliances in similar multi-appliance use categories suggests that concurrent use is much harder to differentiate than single use. The high values of diagonal dominance in the confusion matrix, however, confirm that the proposed classification approach is able to reliably recognize the monitored appliance combinations.

D. Feature Importance Analysis

The proposed system relies on 50 features. These features come from the characteristics of the composite load. A feature importance analysis identified the features that contribute the most to the classification of appliance-combinations.

The most relevant features are PowerPerCurrent, CurrentPerPower, PowerConsistency, PowerRatio, CurrentToPowerRatio, PowerToApparentRatio, ApparentToRealRatio, ReactiveFraction, ReactiveToActiveRatio and ReactivePower.

Out of all features under consideration, the PowerPerCurrent has the highest combined importance of 0.057858 followed by the CurrentPerPower in second with a combined importance of 0.050165 and the PowerConsistency in the third position with a combined importance of 0.047444. Other high importance features were PowerRatio with a combined importance of 0.046649, CurrentToPowerRatio with a combined importance of 0.045712, and PowerToApparentRatio with a combined importance of

0.043946. The results show that classifying appliances can't be done by looking at active power. Relationships between current, power, apparent power, reactive power and power factor offer additional insight into distinguishing appliance operating conditions. The top 15 engineered features and their corresponding importance are presented in Fig. 6.

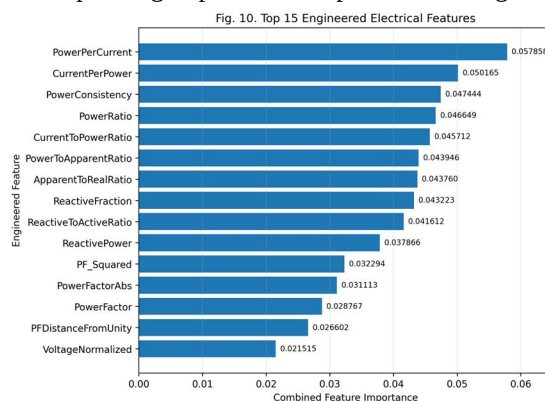


Fig. 6. Fifteen engineered electrical features, ranked by importance, for classifying appliance combinations.

The feature importance results demonstrate the value of engineered electrical characteristics for non-intrusive appliance monitoring since the models can utilize multiple properties of the aggregate electrical signal as opposed to a single parameter.

E. Individual Appliance Detection Performance

In addition to the appliance-combination level classification task, individual device detection performance was also evaluated for the three monitored devices: Phone Charger, Laptop Charger and Hot Glue Gun. The device-wise results are reported in Table IV.

TABLE IV. DEVICE-WISE MULTILABEL CLASSIFICATION PERFORMANCE

Appliance	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Phone Charger	99.40	99.73	99.23	99.48
Laptop Charger	99.99	99.97	100.00	99.99
Hot Glue Gun	99.97	99.93	100.00	99.97

The device-specific models are better, for all three appliances. They give high accuracy. For the Laptop Charger the accuracy is 99.99%. They also give high recall. For the Laptop Charger the recall is 100%. For the Hot Glue Gun the recall is 100% well. The Phone Charger had 99.40% accuracy, 99.73% precision, 99.23% recall, and 99.48% F1-score. Each of these metrics is lower than the best performance for the other appliances, but still very good numbers for individual-level detection.

The exact device combination accuracy was 99.37% for the multi-label outputs. Meanwhile, macro precision, macro recall, and macro F1 were 99.88%, 99.74%, and 99.81% respectively.

This means that the individual-device classification stage provides additional appliance-level information beyond what the joint analysis of eight classes can provide.

F. Appliance-level power estimation

The system also estimates the contribution of individual appliances from the aggregate electrical measurements at the single sensing point.

The learned single-device reference powers used by the system were of 9.00 W for the Phone Charger, 49.90 W for the Laptop Charger, and 9.60 W for the Hot Glue Gun.

When a single appliance is detected, the measured aggregate power can be associated with the corresponding appliance. During simultaneous appliance operation, the system estimates the contribution of each detected appliance utilizing the learned appliance characteristics and adjusts the estimated values with respect to the measured aggregate power.

The results shown are indicative of the system's ability to report the estimated power of the Phone Charger, Laptop Charger and Hot Glue Gun, along with the estimated total power. The estimated contributions from appliances are meant to provide appliance-level information, while only using a single aggregate sensing point. The appliance-level power estimation results are illustrated in Fig. 7

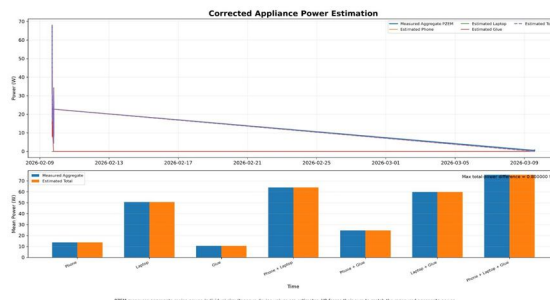


Fig.7. Measured aggregate power and estimated individual appliance power under different appliance operating conditions. An important limitation is that the individual appliance power values are model-based estimates and not direct measurements. Contrary to what someone might think the PZEM-004T measures the power at the sensing point not the power of each individual appliance. Therefore the estimated values, for appliances are not the true values of the individual appliances. The estimated total power is made to equal the measured total power thus making the estimation physical with respect to the electricity meter.

G. Real-Time Monitoring and Reporting

The system we built has a real-time monitoring screen. It shows the readings and details about each appliance. These come from the sensors, in the system. The monitoring interface provides the user with information of aggregate power, individual appliance power, appliance operating status, operating duration, energy consumption, voltage, current, power factor, frequency, detected appliance state, prediction confidence, estimated device power, and estimated electricity cost.

Real-time monitoring dashboard shows measurements and tracks appliance states. The real-time monitoring dashboard also shows estimated power, for appliances displays energy consumption and tells me the electricity cost.



Fig.8 shows the real-time monitoring dashboard of the proposed system.

The system also contains interval reports of time stamped measurements and appliance predictions. The generated 5-minutes report below is just an example of this feature. For the selected 5-minutes interval, the system has recorded 0.004000 kWh in terms of energy consumption with an estimated electricity cost of ₹0.03 while using the displayed tariff of ₹8.00/kWh.

The report also consists of timestamped voltage, current, PZEM power, frequency, power factor detected appliance combination and the estimated appliance power value. An example of the generated monitoring report is shown in Fig. 9.

2026-10-08 10:07:05	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:06	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:07	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:08	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:09	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:10	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:11	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:12	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:13	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:14	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:15	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:16	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:17	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:18	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:19	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:20	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:21	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:22	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:23	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:24	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:25	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:26	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:27	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:28	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:29	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:30	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:31	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:32	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00

Time	V	A	PZEM W	Fz	PF	Detected	Phone W	Laptop W	Gun W	Estimated Total W
2026-10-08 10:07:33	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:34	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:35	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:36	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:37	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:38	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:39	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:40	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:41	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:42	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:43	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:44	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:45	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:46	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:47	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:48	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:49	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:50	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:51	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:52	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:53	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:54	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:55	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:56	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:57	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:07:58	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:00	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:01	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:02	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:03	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:04	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:05	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:06	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:07	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:08	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:09	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:10	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:11	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:12	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:13	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00
2026-10-08 10:08:14	0.00	0.00	0.00	0.00	0.00	AC ON	0.00	0.00	0.00	0.00

Fig.9. Generated monitoring report demonstrating the timestamped electrical measurements, appliance detection, estimated appliance power, energy consumption, and electricity cost.

The reporting feature presents a useful option of examining both the electrical measurements and appliance-specific estimates over the chosen monitoring periods.

```

=====
241.90 V | Current=0.0790 A | Power=8.50 W | Frequency=50.00 Hz | PF=0.440
0.003000 kWh

ENSEMBLE]
on : Phone Charger
ce : 96.51%
: Phone Charger + Laptop Charger (2.79%)

combination: Phone Charger
source : V9 Joint Ensemble
prediction : Phone Charger
prediction : Phone Charger

QUAL DEVICES]
charger ON | prob= 98.00% | threshold=0.620 | power=8.50 W
charger OFF | prob= 0.47% | threshold=0.200 | power=0.00 W
Gun OFF | prob= 3.48% | threshold=0.300 | power=0.00 W

PROBABILITIES]
charger 96.51%
charger + Laptop Charger 2.79%
charger + Hot Glue Gun 0.41%
charger 0.29%
Gun 0.00%
charger + Hot Glue Gun 0.00%
charger + Laptop Charger + Hot Glue Gun 0.00%

power : 8.50 W
d device power: 8.50 W
rgy : 0.003000 kWh

SAFETY]
check : NO

5.224 - - [07/Oct/2026 20:31:59] "POST /predict HTTP/1.0" 200 -

```

Fig.10. Backend prediction output showing realtime electrical measurements, ensemble appliance classification, prediction confidence and estimated device power

H. Discussion

The experimental results demonstrate that the proposed ML-based smart appliance energy monitoring system can identify appliance operating conditions from aggregate electrical measurements with high classification performance.

Among the individual classifiers, the one that demonstrated the best result is HistGradientBoosting with the accuracy of 99.58 % and the macro F1-score of 99.46%. The proposed weighted ensemble achieved an accuracy of 99.55%. Moreover, the macro precision, recall, and F1-score are 99.41%, 99.48%, and 99.44%, respectively. It can be concluded that the weighted combination of HistGradientBoosting, ExtraTrees, and Random Forest classifiers allows achieving high accuracy in classification of the given appliance combinations.

The class-wise results show that individual appliance states can generally be identified with very high reliability. The relatively low recall obtained for the Laptop Charger + Hot Glue Gun and Phone Charger + Laptop Charger + Hot Glue Gun suggests simultaneous appliance operation is a still relatively challenging problem. It should be expected, that within the context of non-intrusive load monitoring, concurrently operated appliances would contribute a combined electrical signature to a composite signature of the measured power signal.

The feature-importance analysis further reveals that power-, current-, apparent-power-, reactive-power- and power-factor-based characteristics were significantly informative for the classification task, suggesting that engineered electrical features, rather than the instantaneous active-power measurement alone, should be employed.

The individual-device evaluation was successful as demonstrated by the high accuracy rates for all three appliances of over 99%. The combination of joint appliance classification and individual-device detection therefore provides both an overall appliance-state prediction and additional appliance-level information.

The final dashboards' tab and interval tabs' results represent the implemented software tool's ability. Using this tool, one can predict the measured electricity values, the appliance's status, the calculated appliance power, energy, and cost with a high degree of accuracy.

However, the results must be used with all considered appliances and data in mind. In particular, individual appliance power values are estimated from aggregate measurements and are not direct measurements of individual appliances. Besides, the last arbitration stage specific to laptops decreased the hold-out classification accuracy from 99.55% to 99.34% in the given assessment. Therefore arbitration should be seen as a method that comes after processing instead of a general way to make things more accurate.

Overall, the results establish that the combination of single-point electrical sensing, 50 engineered electrical features, classical ensemble machine learning, individual-device detection, appliance-level power estimation, and real-time monitoring presented here is an effective lightweight approach for non-intrusive smart appliance energy monitoring.

V. CONCLUSION AND FUTURE SCOPE

An innovative ML-Based Smart Appliance Energy Monitoring System is proposed that is non-intrusive in terms of using the power aggregate measurements at one point to extract the operating conditions of the appliances. The ESP32 acquisition module, PZEM-004T sensing module, Wi-Fi communication, engineered electrical features and machine learning models are utilized for the appliance-level monitoring of the connected loads without using any dedicated sensor for the appliances.

In total, 50 engineered electrical features were considered for the eight-class appliance-combination classification problem. The highest accuracy was achieved by the HistGradient boosting classifier among the individual classifiers at 99.58%. On the contrary, the proposed weighted ensemble model of HistGradient boosting, Extratrees and Random forest achieved 99.55% accuracy with 99.41% macro precision, 99.48% macro recall and 99.44% macro F1-score for all classes. On an individual basis, the accuracies achieved were as high as 99.40%, 99.99%, and 99.97% for the phone charger, laptop charger and hot glue gun respectively.

The achieved results confirm that engineered electrical characteristics, such as the features related to power-current relationship, power consistency, power ratios, and reactive-power related features, carry useful information for the differentiation of the operating conditions of different appliances. The proposed system also provides an estimate of the appliances' power, energy consumption, electricity cost, as well as real-time monitoring and interval-based reporting via the implemented dashboard and reporting interface. However, it should be noted that the obtained individual appliance power values at simultaneous operation of multiple appliances are based on the system's mathematical model using the single-point aggregate electrical consumption. This explains the relatively higher complexity of simultaneous appliance combinations compared with a single appliance, reflected in the achieved lower recall for the multiple appliance operating classes.

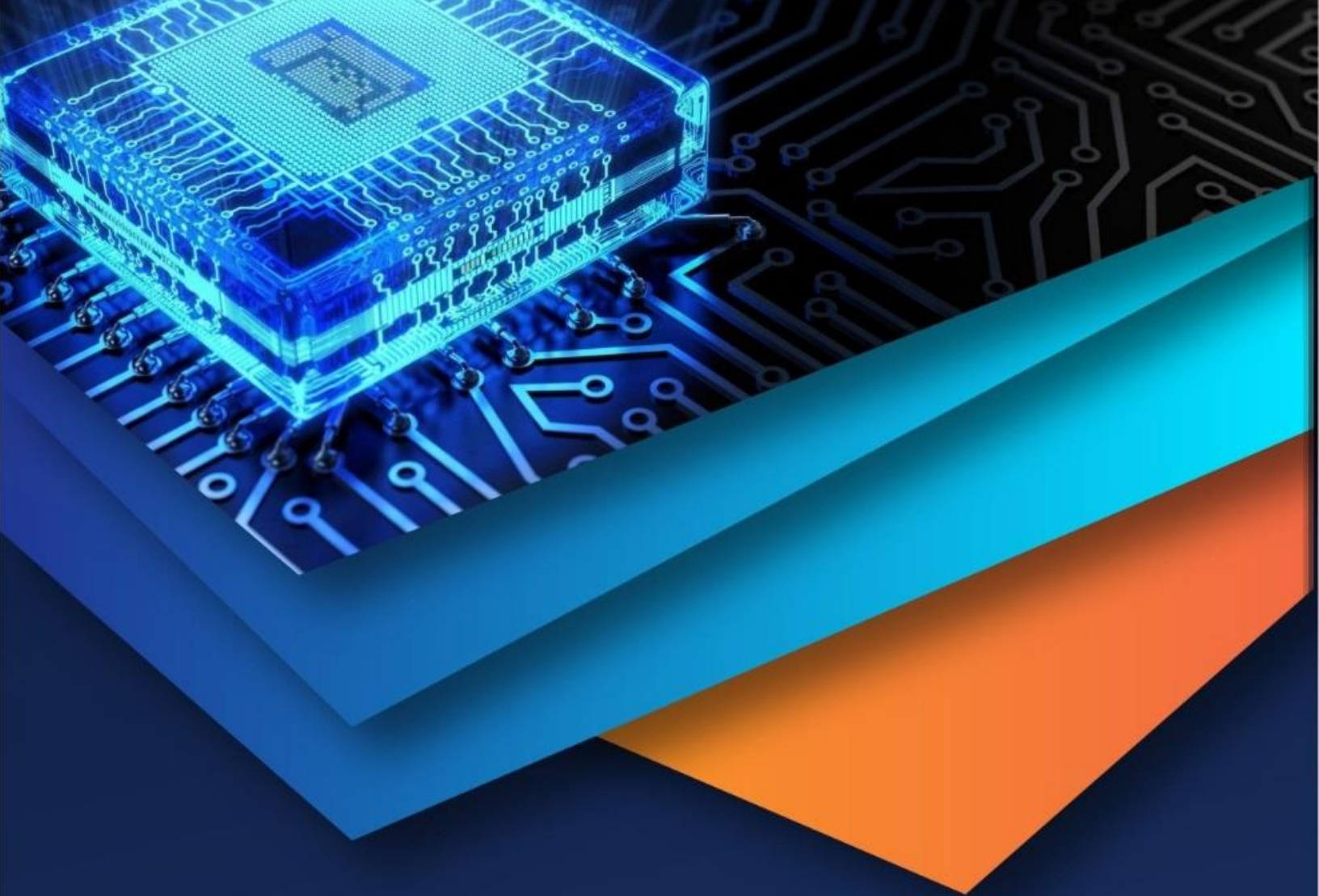
Future work would consider the increase in the number of appliances under monitoring as well as the consideration of higher-resolution current or voltage waveforms to exploit possible transient content in order to further improve the performance. Alternatively, deep-learning-based approaches could be considered if sufficient labeled data becomes available for training such models. The proposed system could also be utilized for intelligent demand-side energy management and appliance scheduling solutions to enable smarter electricity consumption patterns. In summary, the developed solution demonstrated a certain degree of promise in its capacity to allow for smart monitoring of appliance energy by means of single point electrical sensing plus lightweight ensemble machine learning.

VI. ACKNOWLEDGEMENT

The authors wish to express their sincere gratitude to Ms. Priya Meshram, Department of Computer Science and Engineering, G.H Raisoni University, for her valuable guidance, suggestions and support during the entire course of this work.

REFERENCES

- [1] "Machine learning models for non-intrusive load monitoring: A systematic review and meta-analysis," MDPI, Mar. 2026.
- [2] "Lightweight and scalable transfer learning framework for load disaggregation," arXiv:2603.04998, 2026.
- [3] X. Kong, Z. Gui, M. Wu, C. Miao, and Z. Luo, "A hybrid deep learning framework for non-intrusive load monitoring," *Electronics*, vol. 15, no. 2, p. 453, Jan. 2026.
- [4] "Performance enhancement of non-intrusive load monitoring based on adaptive multi-scale attention integration module," *Electronics*, vol. 15, no. 3, Jan. 2026.
- [5] "NILMFormer: Non-intrusive load monitoring that accounts for non-stationarity," arXiv:2506.05880, 2025.
- [6] J. Kelly and W. Knottenbelt, "The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes," *Scientific Data*, 2015.
- [7] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, 1992.
- [8] NILMTK: An open-source toolkit for Non-Intrusive Load Monitoring research.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)