IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 17, 2021, accepted August 9, 2021, date of publication August 11, 2021, date of current version August 24, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3104263

Real-Time Non-Intrusive Electrical Load Classification Over IoT Using Machine Learning

MD. TANVIR AHAMMED[1], MD. MEHEDI HASAN[1], MD. SHAMSUL AREFIN[2], (Member, IEEE), MD. RAFIQUL ISLAM[3], (Senior Member, IEEE), MD. AMINUR RAHMAN[3], EKLAS HOSSAIN[4], (Senior Member, IEEE), AND MD. TANVIR HASAN[1], (Senior Member, IEEE)

[1]Department of Electrical and Electronic Engineering, Jashore University of Science and Technology (JUST), Jashore 7408, Bangladesh
[2]Department of Electrical and Electronic Engineering, Bangladesh University of Business and Technology (BUBT), Dhaka 1216, Bangladesh
[3]Department of Electrical and Electronic Engineering, Khulna University of Engineering and Technology (KUET), Khulna 9203, Bangladesh
[4]Oregon Renewable Energy Center (OREC), Department of Electrical Engineering and Renewable Energy, Oregon Institute of Technology, Klamath Falls, OR 97601, USA

Corresponding author: Md. Tanvir Hasan (tan_vir_bd@yahoo.com)

This work was supported by Jashore University of Science and Technology (JUST), Bangladesh, under the project entitled "Energy Saver System for Domestic Loads".

ABSTRACT In this era of technological advancement, the flow of an enormous amount of information has become such an inevitable phenomenon that makes a path for the takeover of the internet of things (IoT) based smart grid from the currently available grid system. In a smart grid, demand-side management plays a crucial role in reducing the generation capacity by shifting the user energy consumption from peak period to off-peak period, which requires detailed knowledge of the user consumption at the individual appliance level. Non-intrusive load monitoring (NILM) provides an exceptionally low-cost solution for determining individual appliance levels using a single-point measurement. This paper proposed an IoT-based real-time non-intrusive load classification (RT-NILC) system considering the variability of supply voltage using low-frequency data. Due to the unavailability of smart meters at the household level in Bangladesh, a data-acquisition system (DAS) is developed. The DAS is capable of measuring and storing rms voltage, rms current, active power, and power factor data at a sampling rate of 1 Hz. These data are processed to train different multilabel classification models. The best-performed classification model has been selected and utilized for the implementation of RT-NILC over IoT. The Firebase real-time online database is considered for data storage to flow the data in two-way between end-user and service provider (energy distributor). The GPRS module is used for wireless data transmission as a Wi-Fi network may not be available everywhere. Windows and web applications are developed for data visualization. The proposed system has been validated in real-time, using rms voltage, rms current, and active power measurements at a real house. Even under supply voltage variability, the performance evaluation of the RT-NILC system has shown an average classification accuracy of more than 94%. Good classification accuracy and the overall operation of the IoT-based information exchange systems ensure the proposed system's applicability for efficient energy management.

INDEX TERMS Non-intrusive load monitoring, real-time load classification, IoT framework, machine learning, variation of supply voltage.

I. INTRODUCTION

Nowadays, people are more involved in ground-breaking technological research in pertinent areas, especially smart cities, smart homes, the internet of things (IoT), and others. These technological advancements are quite demandable in present times - from the original constituents of a city, smart homes, and factories in the fourth industrial revolution smart cities, which are developed by combining IoT with artificial intelligence (AI). The rapid development in

The associate editor coordinating the review of this manuscript and approving it for publication was Tyson Brooks.

building construction and urbanization increases power demand, and these changes require an efficient energy management program (EEMP), especially for developing countries. EEMP can be obtained by monitoring electric appliances' energy consumption patterns in real-time. To achieve efficient energy management, electric loads must be identified within a given period in real-time from the household or workplace. This load identification process is called load monitoring [1], which can be achieved in intrusive and non-intrusive ways. Non-intrusive load monitoring (NILM) is a single sensor-based process, where the sensor is installed at the power entry point and can recognize turned-on appliances without the requirement of multiple sensors for each appliance. The NILM, an exceptionally cost-effective and promising alternative to intrusive load monitoring, was originally developed by Hart *et al.* in the early 1980s [1]. Recent days extensive research has been done to make this method more reliable and applicable in real-life and real-time load disaggregation. A detailed comparative discussion about NILM is available in Refs. 1 and 2. García-Pérez *et al.* reported that NILM could help energy management of both residential and non-residential buildings [3].

The NILM using high frequency (in the range of kHz and higher) has been observed to attain excellent performance in many studies. Xin *et al.* showed that the frequency domain characteristics had been extracted by analyzing steady-state current to realize load operation states [4]. Ruyi *et al.* utilized 1-16th harmonics and a neural network to differentiate multiple appliances [5]. Le *et al.* proposed the lower odd harmonics-based household appliance classification using a bagging decision tree [6]. A time-frequency analysis-based NILM was also reported by Lin and Tsai [7]. The particle filtering was demonstrated for load disaggregation by Egarter *et al.* [8]. Liu *et al.* reported that the on-off transient signature curve could be employed to identify the appliance [9]. Though the mentioned approaches' performances are satisfying [4]–[7], extra hardware is essential to make the high sampling data collection possible [10].

With the advent of smart grid, smart energy meter is becoming an inseparable part of the energy management systems (EMS). Consequently, NILM with low-frequency data has been the focus of many researchers due to the availability of low-frequency data in smart energy meters. Aiad *et al.* and Cominola *et al.* tried to decompose low-frequency total powers using hidden Markov model-based methods [11], [12]. However, the increasing computational complexity associated with the growing number of appliances makes the methods difficult to implement for real-time applications. Individual appliance identification using features of the low-frequency power-series signal was demonstrated by Corrêa and Castro [13] and Zhang *et al.* [14]. Rafiq *et al.* also identified a single appliance using low-frequency active power (P), apparent power (S), reactive power (Q), rms voltage (V), rms current (I), and power factor (PF) data [15]. Le *et al.* showed a NILM system to identify multiple appliances using transient features from 15 Hz power-series signals by decision tree algorithm [16]. Dinesh *et al.* also demonstrated a NILM system for identifying multiple appliances using low-frequency power-series signals without showing the possibility of determining a single appliance [17]. For the classification of different combinations of multiple appliances, the naïve Bayesian estimation model had been employed by Yang *et al.* [18], while the identification of appliances with a similar power profile was challenging.

For NILM, S, P, and Q have been the most utilized features [19]–[23]. The P and V have recently been observed to obtain high performance even at low frequency [24]. The P, V, and S, I are extracted from AMPds and REDD datasets [25], [26]. Different statistical features, such as Interquartile Range, Crest Factor, Variance, Kurtosis, Mean Absolute Deviation, Skewness, and Form Factor, are extracted from the current waveform and envelope current waveform with the aim of disaggregating loads in [24]. Shapelet extracted from the envelope of the current waveform is used in [26]. Due to high-level redundancy, time-series data is transformed into the frequency domain to extract harmonics information with Fourier transform [27]. Discrete Wavelet Transform (DWT) [28], [29], Stock well Transform [7], and harmonic current-based features [30] have also been studied by researchers.

Both supervised and unsupervised machine learning algorithms have been applied to load classification problems. The model must be trained offline to learn from available data before the supervised method's actual classification. Most common supervised methods include Support Vector Machines (SVM) [26], [29], K-Nearest Neighbors (k-NN) [26], naïve Bayes classifiers [27], Multilayer Perceptron (MLP) [31], Convolutional Neural Networks (CNNs) [32], [33], Deep Neural Networks [34], and Particle-Swarm-Optimization [35].

Welikala *et al.* proposed a power decomposition-based real-time load monitoring system considering voltage variability within the power line [36]. Although they achieved good load identification accuracy, their approach was based on a local machine. One of the most crucial IoT-based EMS issues using NILM is data transmission from localized creation points to the cloud database for further processing and applications. This transmission can be attained with a wide range of technologies and protocols, such as Wi-Fi, GPRS, ZigBee, Bluetooth, etc. [37]–[39]. The smart edge analytics-empowered power meter prototype employs ThingSpeak as cloud storage where only two appliances were considered, such as electric fan and hairdryer for load identification [40]. Moreover, the ethernet connection was considered for data transmission, which is not feasible at all. The IoT-based intrusive load monitoring (ILM) had been proposed by Franco *et al.* for activity recognition in smart homes [41]. As ILM requires a lot of sensors, it is not a cost-effective solution. Most of the previous NILM reports are conducted on local machines using either an online database [13]–[35], [40] or collecting data from smart meters [36]. If NILM can be implemented to observe the

current load condition over the internet, the power authority can observe the current load condition, leading to successfully achieving the EEMP. To the best of our knowledge, there is no report on real-time load monitoring feasibility over IoT considering supply voltage variability in NILM algorithms. Further, none of the studies provides a complete demonstration of a practically feasible end-to-end solution for load monitoring in real-time over IoT. Therefore, a complete practicable end-to-end solution for load monitoring in real-time over IoT is essential. This work demonstrates a practically implementable software and hardware package for real-time non-intrusive load classification (RT-NILC) over IoT using machine learning. The proposed system has been developed by designing a data acquisition system (DAS) that can measure and store rms voltage, rms current, active power, and power factor data to a micro-SD card. These data are employed to train and select a machine learning model for the implementation of RT-NILC. Finally, using some IoT framework, real-time load monitoring over the internet is achieved.

Therefore, the significant contributions of this work are as follows.

1. Designing a low-cost data acquisition system in case of the unavailability of a smart meter.
2. Collection of data using a customized system at different supply voltage from a household in a realistic scenario.
3. Introduction to a novel current decomposition-based on state appliances database building using individual appliance database by combination.
4. Implementation of a real-time IoT-based NILM system using low-frequency data with a complete hardware and software solution.

The rest of the manuscript is organized as follows. The basic overview of the proposed RT-NILC over IoT is discussed in Section II. Section III clearly explains the hardware description of the data acquisition system (DAS). Section IV describes the data preparation procedure for machine learning. Evaluation and selection of machine learning algorithms are discussed in Section V. Section VI explains the complete implementation of RT-NILC over IoT. Results and discussion are included in Section VII. Finally, Section VIII concludes the article.

II. OVERVIEW OF PROPOSED SYSTEM

The proposed RT-NILC system for real-time non-intrusive load monitoring over IoT is implemented utilizing custom-designed hardware, software, and webpage, as shown in **Fig. 1**. The whole work is divided into two stages. At Stage 1, a data acquisition system (DAS) is developed to prepare a database to train different machine learning classification algorithms. Then best machine learning model is selected based on performance scores. The description of DAS and database preparation are described in Sections III and IV, respectively. In Stage 2, the best load classification model has been employed to implement RT-NILC over IoT. The description of RT-NILC is given in Section VI. The hardware used in RT-NILC is an AC meter that reads rms voltage, rms current, and active power data from a house and sends these data to a real-time cloud database using GPRS communication. The software section reads data from the cloud database, makes load classification using the best machine learning model, and sends the classification results to the cloud database. A webpage is hosted on a website to view the current load conditions of a specific house remotely. The software and load classification models were developed in Python 3.8. Firebase is used here as a cloud database, and the website is hosted on 000webhost.com server.

III. DESCRIPTION OF DATA ACQUISITION SYSTEM

Due to the unavailability of smart meters in Bangladesh, a data acquisition system (DAS) has been built using PZEM-004T-100 A [42], Current Transformer (CT), Arduino Uno, micro-SD card module, and RTC module. The required components and their specifications are listed in **TABLE 1**. The detailed description is as follows:

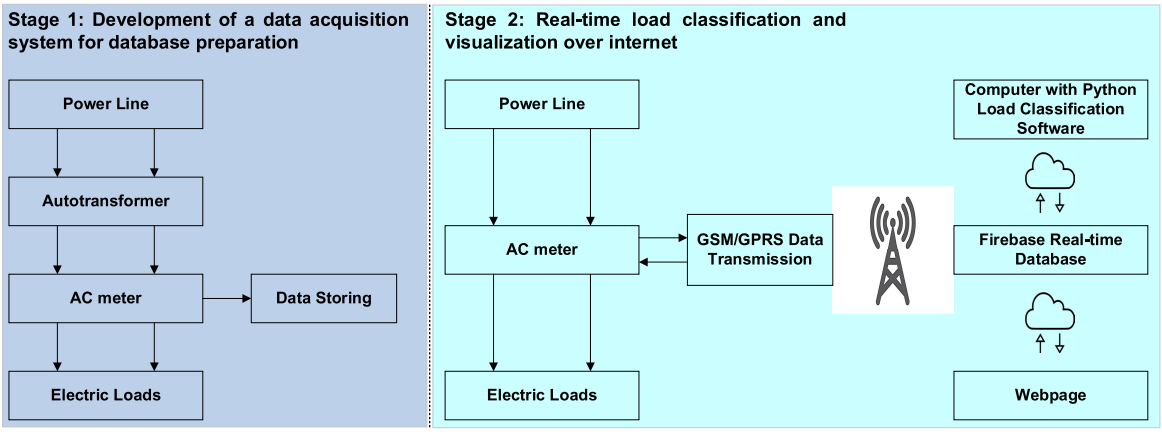


FIGURE 1. Overview of the proposed system.

TABLE 1. Component list for data acquisition system (DAS).

Component name	Specification	Quantity
PZEM-004T-100 A	Voltage measuring range: 80∼260 V, resolution: 0.1 V. Current measuring range: 0∼100 A, resolution: 0.001 A. Power measuring range: 0∼23 kW, resolution: 0.1 W. Power factor measuring range: 0.00∼1.00, resolution: 0.01. Energy measuring range: 0∼9999.99 kWh, resolution: 1Wh. Frequency measuring range:45∼65 Hz, resolution: 0.1 Hz. Working temperature: -20°∼+60°C	01
Current Transformer (CT)	Current measuring range: 0∼100 A Ratio: 1000:1	01
Arduino Uno	Operating voltage: 5 V Flash memory: 32 kB Clock speed: 16 MHz Serial interface: 0(Rx), 1(Tx), TTL	01
Micro SD TF card module	Operating voltage: 4.5-5.5 V, 16 GB	01
DS3231 RTC module	Operating voltage: 5 V Counts Seconds, Minutes, Hours, Date of the Month, Month, Day of the Week, and Year, with Leap-Year Compensation Valid up to 2100	01
DC power supply	Input (AC): 180-240 V, 50 Hz Output (DC): 5 V, 1 A	01

PZEM-004T: It is an AC communication module capable of measuring rms voltage, rms current, active power, frequency, power factor, and energy. It sends data via TTL serial interface, which is compatible with software and hardware serial communication with Arduino.

CT: 100A CT (1000:1) is used to reduce current to PZEM-004 T's compatible level.

Arduino: Arduino Uno is used here to read the TTL serial data from PZEM-004T-100 A and stores the data in comma-separated values (.csv) format in a micro-SD card with date-time information.

Micro SD TF card module: A 16 GB micro-SD card is used to store data from PZEM-004T-100 A.

RTC module: DS3231 RTC module is used here for data acquisition date and time information.

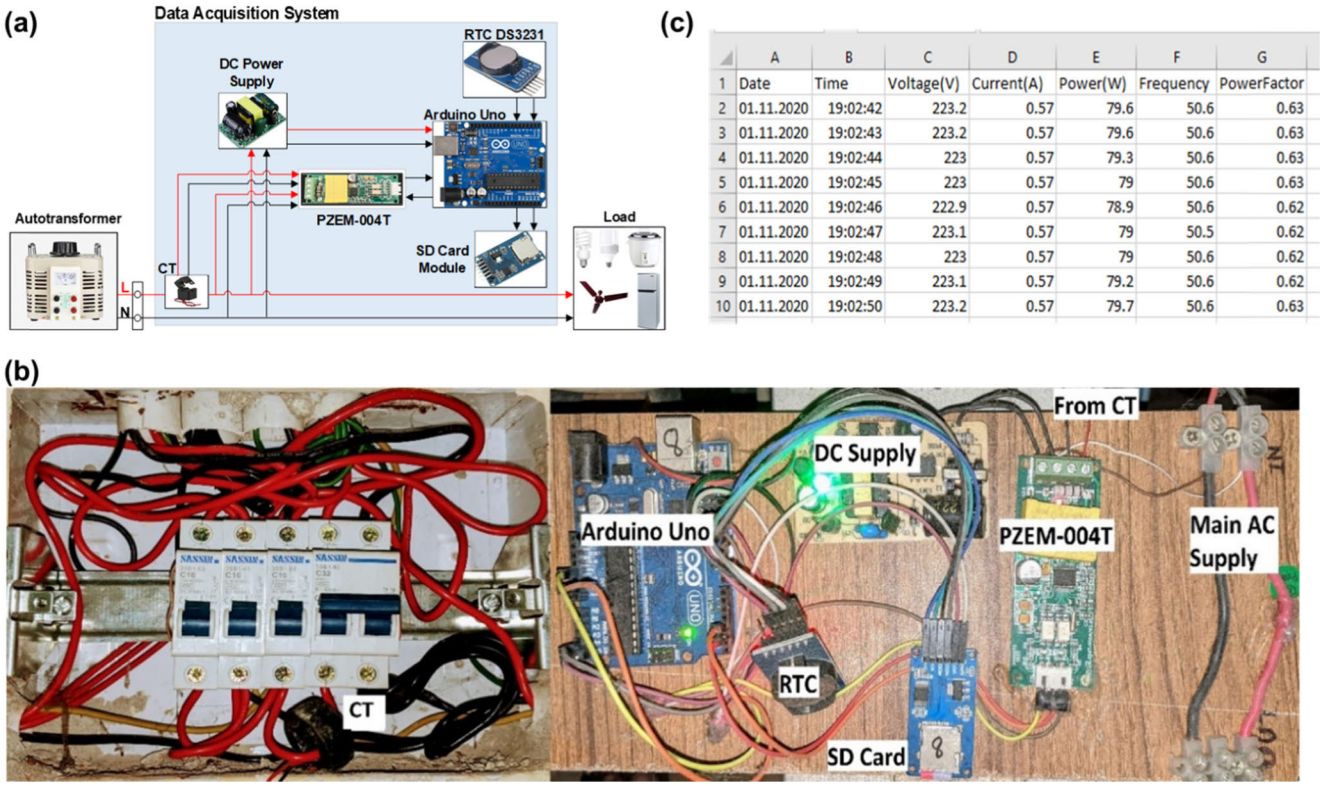
DC power supply unit: A 5 V, 1 A DC power supply is used to power all of the hardware components mentioned above.

The block diagram (a), hardware implementation of DAS (b), and a snapshot of recorded data (c) are shown in **Fig. 2**. **Figure 2 (b)** also shows the CT connection to a distribution board and the DAS. The data has been stored in the micro-SD card using the developed DAS, as shown in **Fig. 2 (c)**. To ensure the DAS's measurement accuracy, it is calibrated with the standard calibration meter named Fluke 5502A [43]. The rms voltage, rms current, active power, and power factor are measured for four types of appliances by varying the supply voltage from 200 to 230 V using the developed DAS and standard Fluke 5502A meter. TABLE 2 shows the comparative measurement data and % error in terms of power. It is found that there is a slight difference between the measurement data of DAS and the standard Fluke 5502A meter. Thus, it can be concluded that the DAS has a good measurement accuracy. The total cost of the DAS is ∼$ 22, which is very low compared to the commercially available system [44].

IV. PREPARATION OF DATABASE

One-minute data from six commonly used household appliances (rice cooker, LED lamp, CFL lamp, water heater, fridge, and ceiling fan) are collected using the DAS. **TABLE 3** shows the rated wattage ratings of the six appliances used here. The sampling frequency of the data acquisition system is 1 Hz (1 data/second). To incorporate the variability of the supply voltage, data from each appliance are collected at different voltages from 210 to 240 V with a 2 V increment per step. This voltage range (210 – 240 V) is selected because the supply voltage to the residence usually remains within this range. The voltages are changed using a single-phase autotransformer (3 kVA, 0 - 250 V, 50 Hz) [45]. Therefore, the data length is $60 \times 16 = 960$ for each appliance for sixteen different voltage levels. During data collection regulator of the ceiling fan was fixed to a particular position. The effect of voltage variability is shown in **Fig. 3**, which shows the change of power consumption of LED, CFL, water heater (WH), and rice cooker (RC) with a change in supply voltage ranging from 221 to 229 V, and it is found that there exists a significant difference in power consumption for resistive loads e.g. RC and WH while a negligible effect for non-resistive loads for instance CFL and LED. For instance, at 221 V, RC and WH's power consumption is 975 W and 678.4 W, respectively. While the power consumption of RC and WH are 1015W and 728 W, respectively, at 229 V. The power consumption difference of 40 W and 49.6 W have been observed for RC and WH, respectively, by changing supply voltages. The combined power difference of RC and WH is 89.6 W, which is between the fridge's power rating (100 W) and the ceiling fan (80 W). Thus, the supply voltage variation may negatively affect the load classification accuracy if it is not accounted for. For instance, at 229 V, only rice cooker may be misclassified as rice cooker plus LED or CFL as LED and CFL alone have power consumption near about 32 W.

The developed DAS is installed at a house, where rms voltage, rms current, active power, and power factor data are collected by turning ON only one appliance at a time for 1-minute. This process has repeated for sixteen different voltage levels (210 - 240 V, 2 V increment per reading), as discussed earlier. At the same time, the data are stored in a micro-SD card. Since individual appliance data are collected, and any combinations of appliances may occur, it is required to calculate the rms current, active power, and power factor data for each possible combination. The goal is to determine all possible ON-state rms current, active power,



	A	B	C	D	E	F	G
1	Date	Time	Voltage(V)	Current(A)	Power(W)	Frequency	PowerFactor
2	01.11.2020	19:02:42	223.2	0.57	79.6	50.6	0.63
3	01.11.2020	19:02:43	223.2	0.57	79.6	50.6	0.63
4	01.11.2020	19:02:44	223	0.57	79.3	50.6	0.63
5	01.11.2020	19:02:45	223	0.57	79	50.6	0.63
6	01.11.2020	19:02:46	222.9	0.57	78.9	50.6	0.62
7	01.11.2020	19:02:47	223.1	0.57	79	50.5	0.62
8	01.11.2020	19:02:48	223	0.57	79	50.6	0.62
9	01.11.2020	19:02:49	223.1	0.57	79.2	50.6	0.62
10	01.11.2020	19:02:50	223.2	0.57	79.7	50.6	0.63

FIGURE 2. Schematic block diagram of the DAS (a), hardware implementation (b) and snapshot of recorded data format (c).

TABLE 2. Measurement accuracy of the developed data acquisition system (DAS).

Load	Developed DAS				Standard meter (Fluke-5502A)				% Error
	RMS Voltage, V (V)	RMS Current, I (A)	Active Power, P (W)	Power Factor, PF	RMS Voltage, V (V)	RMS Current, I (A)	Active Power, P (W)	Power Factor, PF	$(\frac{P_{Fluke}-P_{DAS}}{P_{Fluke}})\times 100$
1000 W, Rice Cooker	200	4.01	802	0.998	200	4.01	802.1	1	-0.01
	209.9	4.214	884.14	0.998	209.8	4.21	884	1	0.02
	219.9	4.413	970.24	0.998	219.9	4.41	970.2	1	0.01
	229.9	4.443	1021.23	0.998	229.7	4.45	1023.5	1	-0.22
30 W, LED	199.9	0.195	30.14	0.757	199.8	0.20	30.3	0.77	-0.53
	210.1	0.175	30.14	0.786	210	0.18	30.2	0.8	-0.20
	219.9	0.166	30.83	0.819	220	0.17	30.8	0.82	0.09
	229.9	0.156	30.52	0.884	229.8	0.15	30.5	0.88	0.07
30 W, CFL	200	0.208	26	0.648	200	0.20	25.9	0.65	0.38
	209.9	0.221	28.65	0.624	210	0.22	28.6	0.62	0.17
	219.9	0.234	30.35	0.598	220	0.23	30.2	0.6	0.49
	229.9	0.244	31.47	0.574	229.9	0.24	31.4	0.57	0.22
80 W, Ceiling Fan	199.9	0.358	57.60	0.798	199.91	0.36	57.6	0.80	0
	210	0.381	67.13	0.843	209.97	0.38	67.1	0.84	-0.04
	219.9	0.398	76.56	0.867	219.93	0.40	77.1	0.88	-0.70
	229.8	0.427	89.11	0.898	229.89	0.43	89.2	0.90	-0.10

and power factor data of all appliances for each voltage. Therefore, the prepared database should contain the combined rms current, combined active power, combined power factor, and the name of each combination's appliances for each voltage level. In each combination, rms current of the individual appliance is decomposed into the rectangular form using equations (1-4).

$$I_{max} = \sqrt{2}I \tag{1}$$

$$\emptyset = \cos^{-1}(PF) \tag{2}$$

$$I_{RE} = I_{max}\cos\emptyset \tag{3}$$

$$I_{IM} = I_{max}\sin\emptyset \tag{4}$$

TABLE 3. Ratings of employed appliances.

Appliance Name	Rated Wattage (W)
Rice cooker	1000
LED lamp	30
CFL lamp	30
Water heater	700
Fridge	100
Ceiling Fan	80

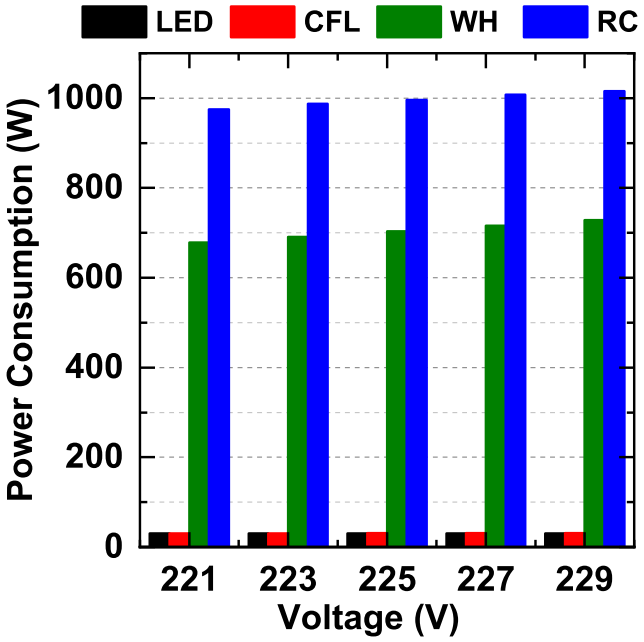


FIGURE 3. Power consumption of LED, CFL, rice cooker, and water heater with voltage change.

TABLE 4. Possible combinations of appliances with the corresponding label.

Label	Load combination	Label	Load combination	Label	Load combination
0	L	21	L+C+RC	42	L+C+RC+F
1	C	22	L+C+WH	43	L+C+RC+FR
2	RC	23	L+C+F	44	L+C+WH+F
3	WH	24	L+C+FR	45	L+C+WH+FR
4	F	25	L+RC+WH	46	L+C+F+FR
5	FR	26	L+RC+F	47	L+RC+WH+F
6	L+C	27	L+RC+FR	48	L+RC+WH+FR
7	L+RC	28	L+WH+F	49	L+RC+F+FR
8	L+WH	29	L+WH+FR	50	L+WH+F+FR
9	L+F	30	L+F+FR	51	C+RC+WH+F
10	L+FR	31	C+RC+WH	52	C+RC+WH+FR
11	C+RC	32	C+RC+F	53	C+RC+F+FR
12	C+WH	33	C+RC+FR	54	C+WH+F+FR
13	C+F	34	C+WH+F	55	RC+WH+F+FR
14	C+FR	35	C+WH+FR	56	L+C+RC+WH+F
15	RC+WH	36	C+F+FR	57	L+C+RC+WH+FR
16	RC+F	37	RC+WH+F	58	L+C+RC+F+FR
17	RC+FR	38	RC+WH+FR	59	L+C+WH+F+FR
18	WH+F	39	RC+F+FR	60	L+RC+WH+F+FR
19	WH+FR	40	WH+F+FR	61	C+RC+WH+F+FR
20	F+FR	41	L+C+RC+WH	62	L+C+RC+WH+F+FR

*L=LED, C=CFL, RC=Rice cooker, WH= Water Heater, F=Fan, FR= Fridge

where, $\emptyset$ is the phase difference between voltage and current, PF represents power factor, I_{RE} represents the real part of current, and I_{IM} is the imaginary part of the current. Hereafter, the real and imaginary parts of current from different appliances have been added in any possible combination to create the total real and imaginary current, I_S component for that particular combination using equation 5. The power factor and the maximum total current I_T for each combination are calculated using equations 6 and 7, respectively.

$$I_S = \sum I_{RE} + j \sum I_{IM} \tag{5}$$

$$Power\ Factor = tan^{-1}\left(\frac{\sum I_{IM}}{\sum I_{RE}}\right) \tag{6}$$

$$I_T = \sqrt{I_{RE}^2 + I_{IM}^2} \tag{7}$$

Finally, the total rms current I_{LC} is obtained for each load combination using equation 8.

$$I_{LC} = \frac{1}{\sqrt{2}} I_T \tag{8}$$

The total active power, P_{TA} of each combination is calculated using a simple summation of active power, P_A of each load in each combination using equation 9.

$$P_{TA} = \sum P_A \tag{9}$$

For N appliances, the possible combination is 2^N-1. Therefore, the total number of possible combinations of six appliances are 2^6-1 = 63. During database preparation, in each combination, the appliance's name is assigned with a unique label (0 - 62). The appliances combination and corresponding label are listed in **TABLE 4**. The 63 combinations of appliances are constructed by adding current vectors and active power data. As a result, the training database consists of 960 × 63 = 60480 data samples of active power (P), rms voltage (V), rms current (I), power factor (PF), and corresponding label. The procedure of database preparation from individual appliance's data is described in **Algorithm 1**. Similarly, two separate databases are prepared for two different voltage levels at 221 and 227 V to verify the performance under a new unknown database which is not included during

Algorithm 1 Training Database Preparation

Input: V, I, P, PF data of N appliances
Output: 2^N-1 data of V, I, P, PF and Label
for $j \leftarrow 1$ **to** N **do**
  data[j] ← **read** V, I, P, PF data for N appliance
/* Store V, I, P, PF data in N different variables */
applianceName[] ← **make** list of N appliances name
for $k \leftarrow 1$ **to length**(data[1]) **do**
  read V[k], I[k], P[k], PF[k] data for N appliance
  v ← V[k] for N appliance
  pf ← PF[k] for N appliance
  i ← **convert** I[k] to rectangular form using PF[k] for N appliance
  p ← P[k] for N appliance
  make list of i, p for N appliances
  I_comb ← **make** combinations of N elements in the list i
  P_comb ← **make** combinations N elements in list the p
  name_comb ← **make** combinations of N elements in list the applianceName
  for $n \leftarrow 1$ **to length**(I_comb) **do**
    I← **abs**(**sum**(I_comb[n])) for each combination
    I_real← **take** real part of I_comb[n] for each combination
    I_imag← **take** imaginary part of I_comb[n] for each combination
    PF← **calculate** power factor using I_real and I_imag for each combination
    P← **sum**(P_comb[n]) for each combination
    V← **round**(v[n]) for each combination
    Label← name_comb[n] of each combination
    trainData← **append** V, I, P, PF and Label data for each combination
return trainData

the training. The new test database consists of $2 \times 60 \times 63 = 7560$ data samples of P, V, I, PF, and corresponding label.

V. EVALUATION AND SELECTION OF MACHINE LEARNING MODEL

The training database is employed to train different supervised machine learning classification algorithms such as Random Forest (RF), XGboost, k-nearest neighbors (KNN), and Naïve Bayes. The detailed information of Random Forest (RF), k-nearest neighbors (KNN), and Naïve Bayes algorithms are available in scikit learn [46] package of Python, while the detailed information of XGboost classification algorithm can be found in Ref. 47. A 10-fold cross-validation has been performed on the prepared train dataset using those classification algorithms for four different sets of features. In this work, the Accuracy, F1-score, Precision, and Recall are considered as classification metrics.

$$\%\ Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (10)$$

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TN}{TP + FN} \quad (12)$$

$$F1\text{-}score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (13)$$

where, TP, FP, TN, and FN are denoted the number of true positive, false positive, true negative, and false negative instances, respectively. The goal is to train the machine learning algorithms to choose the best combination of hyperparameters for each algorithm (RF, XGboost, KNN) using the training dataset. Therefore, hyperparameter tuning (for RF *n_estimators, max_depth, max_features, criterion, min_samples_split, min_impurity_decrease, bootstrap*; for Xgboost *n_estimators, max_depth, min_child_weight, tree_method, eta, gamma, objective*; and for KNN *n_neighbors, weights, algorithm, leaf_size, p*; are used as tuning parameters) is performed by using 10-fold cross-validation (in each iteration 9 subgroups were used to train the model, and the rest one is used for testing the model) using *sklearn.model_selection.Randomized SearchCV* library for each machine learning algorithm except the Naïve Bayes. The Naïve Bayes algorithm is excluded from hyperparameter tuning because it does not have any iterable parameters. Using the library, the accuracy on 10 test datasets (of cross-validation) is computed 250 times for the random combinations of the hyper-parameters for each machine learning algorithm. Hereafter, the best combination of the hyper-parameters, which shows the maximum mean accuracy on cross-validation test datasets, are selected. The hyperparameter tuning is performed on four different sets of features in Python programming language on a single computer (*Intel®Core(TM) i5* - 6200*U* CPU @ 2.30 *GHz*, 8.0 *GB* RAM) with *Windows* - 10 operating system. The outcome results for four different sets of features are shown in **TABLE 5**. It is found that there is a significant improvement in classification score when the rms voltage is incorporated as a feature. It is also observed that the RF and XGboost classifiers represent better performance scores than the others for the feature set V, I, and P. In addition, the RF classification algorithm shows slightly better performance as compared with XGboost. To choose the best classification algorithm, the RF (*n_estimators = 500, max_depth = none, max_features = 'log2', criterion = 'gini', min_samples_split = 2, min_impurity_ decrease = 0, bootstrap = True*) and the XGboost (*n_estimators =60, max_depth = 20, min_child_weight = 1, tree_method = 'auto', eta = 0.3, gamma = 0, objective = 'multi*: *softmax'*) models are stored using the pickle library of Python for further validation. Hereafter, the new unknown test dataset which consists of rms voltage, rms current, active power, and the corresponding label for two different voltage levels

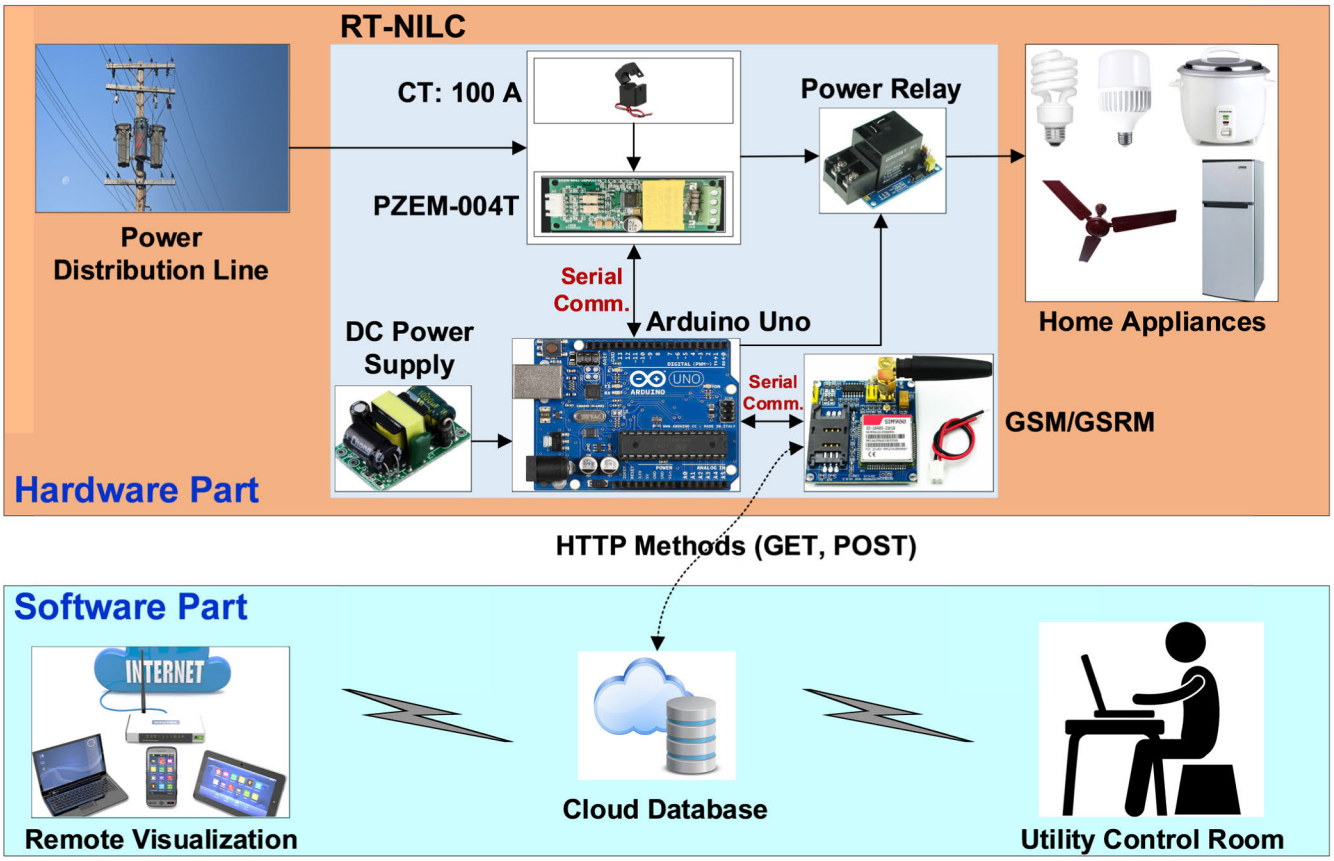


FIGURE 4. Methodology of the proposed RT-NILC over IoT.

TABLE 5. Performance scores using four different sets of features.

Features	10-Fold Cross-Validation Scores (% Maximum Mean Accuracy)			
	Random Forest Classifier	XGBoost Classifier	KNN Classifier	Naïve Bayes Classifier
I, P	89.01	89.40	77.26	83.70
I, P, PF	89.40	89.38	77.24	83.74
V, I, P	95.52	94.43	86.45	83.72
V, I, P, PF	94.10	93.38	86.45	83.76

TABLE 6. Performance scores of random forest and XGBoost classifier using new validation dataset.

Performance scores	Random Forest Classifier		XGBoost Classifier	
	For 221 V (rms)	For 227 V (rms)	For 221 V (rms)	For 227 V (rms)
Accuracy (%)	95.05	94.14	91.23	92.24
F1-score (%)	96.56	96.87	90.28	91.92
Precision (%)	97.10	97.58	91.37	92.96
Recall (%)	96.03	96.18	89.22	90.89

221 and 227 V are considered for the validation. The validation scores of the RF and XGboost classification algorithms are listed in **TABLE 6**. Once more, the RF classifier shows better scores than XGboost. Therefore, the RF classifier has been chosen for the implementation of RT-NILC over IoT.

VI. IMPLEMENTATION OF RT-NILC OVER IoT

To develop the RT-NILC over IoT, it is essential to modify the DAS. A wireless communication medium is required to send the V, I, P data and receive the classification results from a cloud database. The block diagram of hardware components of the proposed RT-NILC is shown in **Fig. 4**. The relay serves an additional purpose like turn ON/OFF the main supply of a house. Since Wi-Fi network may not be widely available, GPRS wireless communication system is considered. For GPRS communication with the cloud database, the SIM900A module, is employed. The workflow diagram and algorithm of the proposed RT-NILC over IoT are shown in **Fig. 4** and **Algorithm 2**, respectively, which has two parts, such as Hardware part and Software part. The proposed system has been developed using an IoT infrastructure that consists of four layers, as shown in **Fig. 5**. The bottom layer, named the device layer, has two sub-layers, such as things and gateway. The measurement sensors (PZEM-004T and CT: 100 A) are the parts of things layer, which measures V, I, P, and PF

data, whereas Arduino Uno and SIM900A are in the gateway layer, building connections between the components of the things layer and network layer. The network layer uses GPRS communication for data transmission from the device layer to the cloud service layer. The cloud service layer consists of a hosting server and Firebase real-time database for data storage and retrieval. The application layer is the top layer that provides service to the end-users.

Algorithm 2 Hardware and Software Part

```
/*                Start of Hardware Part                    */
while Data.available() do
    read V, I, P from a house
/*  using serial communication of PZEM-004T and
    Arduino Uno                                             */
    send V, I, P data to cloud database
/* using SIM900A GPRS communication using http
    GET()  method                                           */
    read OnOff control data from cloud database
/*  using SIM900A GPRS communication using http
    GET() method                                            */
    if  OnOff==0 then
        turn Off main supply
    end if
    if  OnOff==1 then
        turn On main supply
    end if
    delay(seconds)
/*              End of Hardware Part                        */
/*              Start of Software Part                      */
while DataReading.available() do
    read V, I, P from cloud database
    label ← make classification
/*  using V, I, P as feature and previously saved ML
    model                                                   */
    PF ← calculate power factor using V, I, P
    send label to cloud database
    read OnOff control data form user
    send OnOff control data to cloud database
    display V, I, P, PF and name of running appliances
    name
    delay(seconds)
/*                End of Software Part                      */
```

A. HARDWARE PART

This part involves reading data from a house and transmitting data to a cloud database. During reading data, Arduino reads real-time rms voltage, rms current, and active power using serial communication with PZEM-004T. These three data are the features required for the load classification. For transmission of these data to a cloud database, it is required to create a real-time database in a cloud server. Firebase real-time database have been employed to send and receive data. As Arduino Uno cannot generate https connection which is essential to send and receive data in the Firebase database, a PHP proxy server is built in www.000webhostapp.com. The Arduino Uno communicates with the 000webhost PHP proxy server via http protocol (GET(), POST() method) using SIM900A and the PHP proxy server makes https connection (GET(), SET() methods) with Firebase. By this way, it is possible to transfer data between Arduino Uno and Firebase. The hardware is configured to send data to Firebase at 0.2 Hz sampling rate. Here cost-free limited service of Firebase and 000webhost is used for demonstration. **Figure 6** shows the firebase real-time database. Field-1 shows the rms voltage, rms current, and active power, Field-2 represents the classification result, Field-3 represents the ON/OFF status of the main supply to a house. The Arduino sends data to the Field-1 and reads data from the Field-3. Based on the Field-3's value, Arduino turns ON and OFF the main supply to the load (1 = ON, 0 = OFF) using the relay as shown in **Fig. 4**. The Field-2 is associated with the software part.

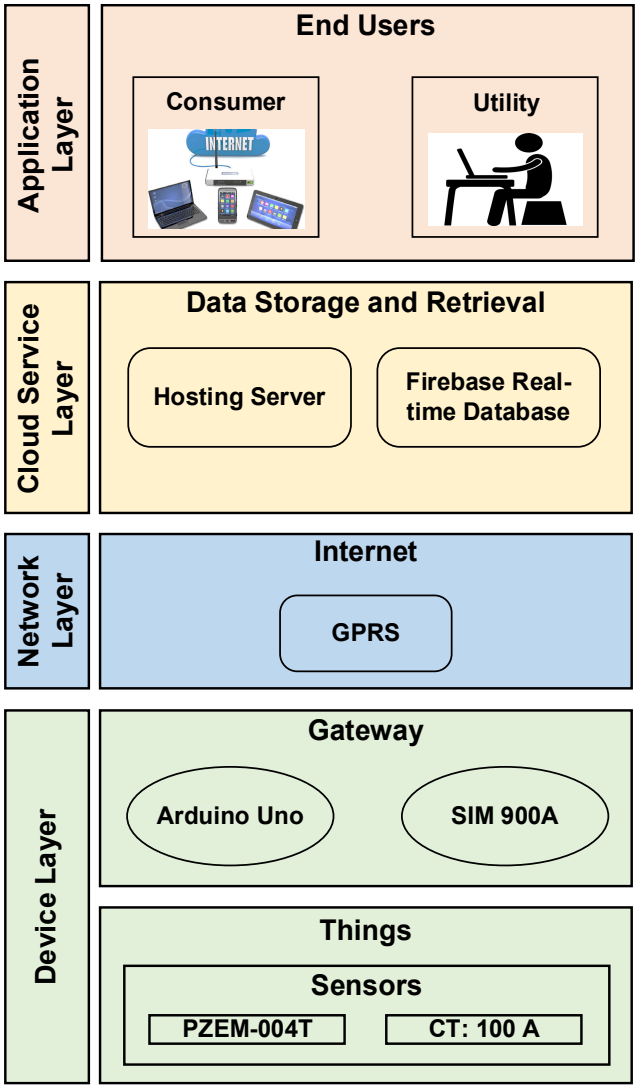


FIGURE 5. Considered IoT architecture consists of four layers.

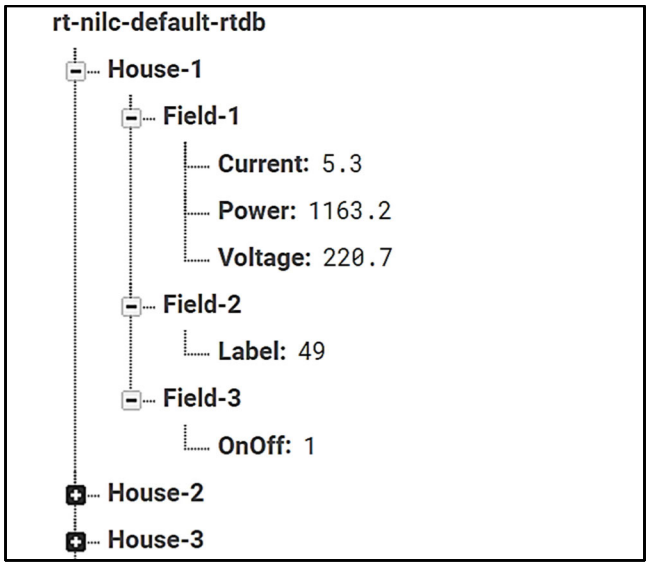


FIGURE 6. Process of firebase real-time database.

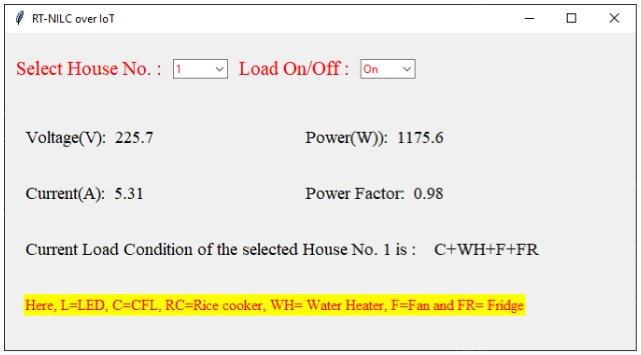


FIGURE 7. Developed GUI layout of RT-NILC.

B. SOFTWARE PART

This part deals with three tasks. Firstly, it involves the extraction of feature data (i.e., rms voltage, rms current, and active power) from Firebase to classify the load combination label using the RF model. Secondly, sending the information of that label to Field-2 of Firebase. Finally, data visualization in a graphical user interface (GUI). To perform the above-mentioned tasks, a desktop GUI application and web application have been developed. The desktop GUI application has been built using tkinter library package of Python and hereafter has been converted to a windows application using pyinstaller library package of Python. The layout of the GUI is represented in **Fig. 7**. The GUI shows the rms voltage, rms current, active power, power factor, and current load status of the selected house. The power factor, which is displayed in the GUI, has been calculated using equation 14.

$$Power\ Factor = \frac{Active\ Power}{Apparent\ Power} = \frac{Active\ Power}{V_{rms} \times I_{rms}} \quad (14)$$

It is noted that the computer which runs the GUI application should have a good internet connection. This GUI application does not require any third-party online machine learning hosting services.

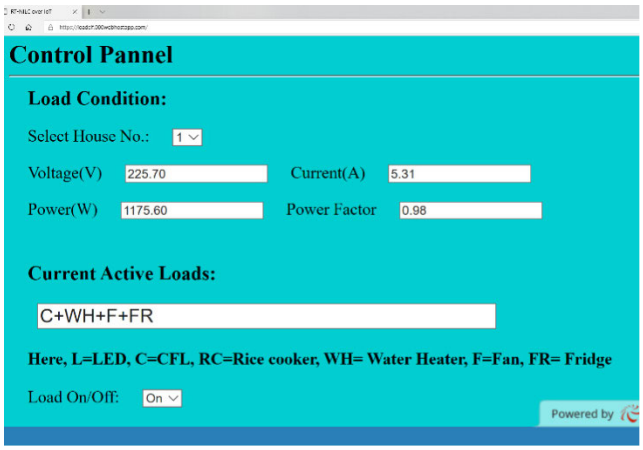


FIGURE 8. Layout of the webpage.

In addition, a webpage is developed using HTML and JavaScript and hosted on 000webhost server. The webpage requires a username and password. By providing correct information, currently running load status, rms voltage, rms current, active power, and power factor information of the selected house where the device is installed can be seen remotely. The power factor data is obtained using Equation 14 in JavaScript. **Figure 8** shows the webpage layout. The desktop application should be in operation to view the running load combinations for a house on the web page. Besides, in the proposed system, the old data sets are replaced by new data sets, and only 699 Bytes of storage is occupied in Firebase database for a user. Thus, a large storage is not required for the proposed system.

VII. RESULTS AND DISCUSSION

This section discusses the experimental results of RT-NILC over IoT, implementation cost with execution time, and a comparative study with the existing work.

A. PERFORMANCE EVALUATION

After complete installation, ten readings have been taken for each of 63 combinations of loads from a real house without regulating the supply voltage, which varies between 220 to 235 V during the experiment. The hardware sends feature data (V, I, P) at 0.2 Hz (in every five seconds) sampling rate to the Firebase. The classification predicted results of RT-NILC over IoT have been taken from the webpage, which are compared with the actual load combinations. The predicted labels and true labels are shown in the confusion matrix in **Fig. 9**. It is observed that the classifier classifies accurately in every label except labels 7, 8, 25, 28, 29, 47, 48, 50, and 60. Label 7 (LED + Rice Cooker) misclassified 2 times out of 10 as label 11 (CFL+ Rice Cooker). The classifier only misclassifies CFL as LED and vice versa. Similar results are found in label 28 (LED + Water heater + Fan), which is misclassified 3 times out of 10 as label 34 (CFL + Water heater+ Fan). Also label 47 (LED + Rice Cooker + Water heater + Fan) is misclassified 5 times out of 10 as label 51 (CFL + Rice Cooker + Water heater + Fan). These

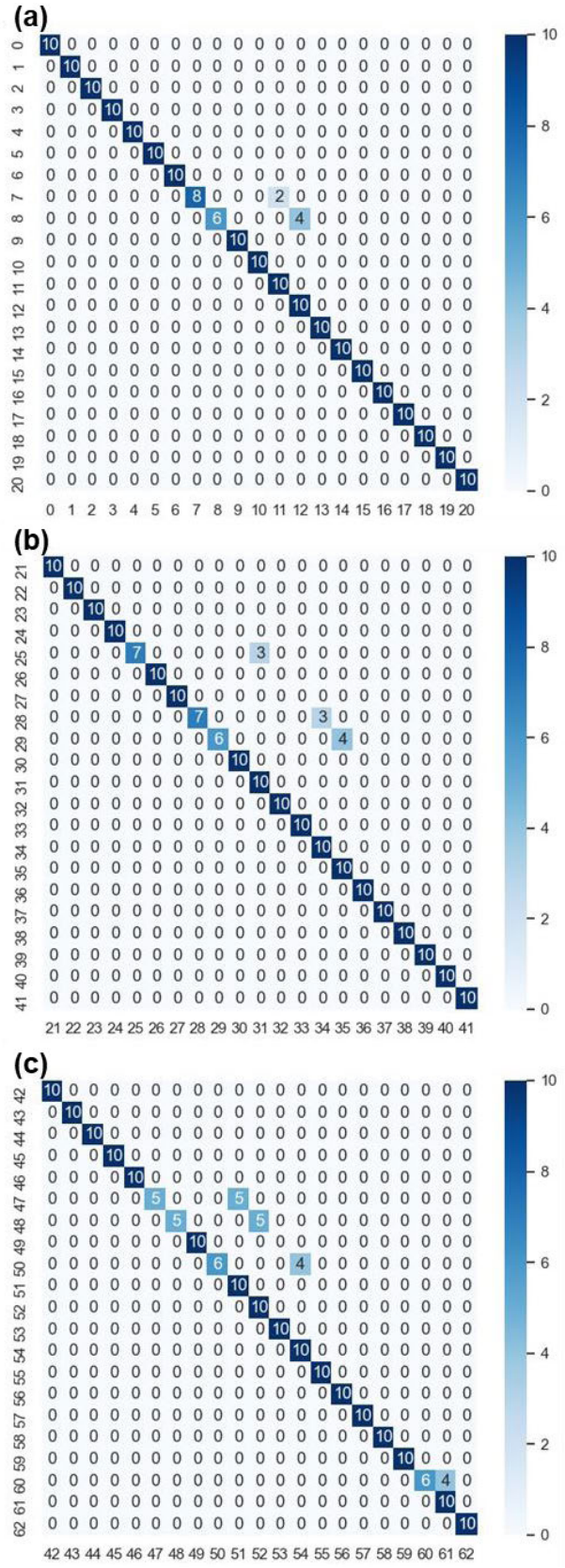


FIGURE 9. Confusion matrix; (a) Confusion matrix for label 0 to 20 (b) Confusion matrix for label 21 to 41 (c) Confusion matrix for label 42 to 62.

TABLE 7. Performance scores of the implemented RT-NILM over IoT.

Performance metrics	Overall scores
Accuracy (%)	94.60
Precision (%)	96.13
Recall (%)	94.60
F1-score (%)	95.36

inaccurate results may be due to the similar wattage ratings of the CFL and LED. The performance scores of the implemented RT- NILC over IoT are calculated using Equations 10-13, then averaged and listed in **TABLE 7**. Even under realistic scenarios, the proposed system has produced very accurate results with an average accuracy of more than 94%.

B. IMPLEMENTATION COST AND EXECUTION TIME

The required hardware components are shown in **Fig. 4**. The cost of the hardware part is ∼$35. The software part requires the purchasing of hosting and data plan for GPRS communication. If the RT-NILC sends V, I, P data every five seconds to the Firebase using SIM900A, then ∼200 Megabytes (MB) of data per month per user is required for the GPRS communication. The per month MB uses can be minimized by increasing data transmission delay in the hardware section. Since the price of web hosting and mobile data plans varies from country to country, it isn't easy to address the overall cost of the proposal. Moreover, the proposed system uses a cloud database and webpage, there might be a concern regarding security and privacy issues. Also, the proposed IoT-based system requires a secure internet connection for data transmission. Proper protocols and robust coding might ensure cybersecurity, whereas the internet connection quality entirely depends on the service providers.

The power consumption curve of the RT-NILC is shown in **Fig. 10** for twenty minutes. The power consumption of the IoT-based RT-NILC system is < 4 W when the relay (shown in **Fig. 4**) is energized, and the rest of the time, it consumes < 3.5 W. When the GUI software (**Fig. 7**) of the RT-NILC system is executed on a computer (*Intel®Core* (*TM*) *i5* - 6200*U* CPU @ 2.30 *GHz*, 8.0 *GB* RAM with *Windows* - 10 operating system), the average time to display a result in the GUI is 3.884 sec (which includes reading feature data (i.e., V, I and P) from Firebase, making classification and sending classification label to Firebase). During performance evaluation, the proposed RT-NILC system is configured to send feature data every five seconds to the Firebase. The proposed system requires < 10 sec to display a new result on the webpage (**Fig. 8**). However, the total execution time can be reduced by decreasing the delay time (which is set to 5 sec in this work). Besides, the data transmission speed from an IoT device to a cloud database depends on the internet connection speed. If a faster communication speed is required, the 4G/3G (such as SIM7600CE shield for Arduino, which

TABLE 8. A comparative study with existing low-frequency NILM systems.

Authors	Sampling frequency used in database preparation	Algorithm	Performance	Real-time implementation over the internet	Purpose
Corrêa et al. (2020) [13]	≤ 1 Hz	Auto-associative Neural Network (AANN) and Multi-Layer Perceptron Neural Network (MLP)	F1-score = 0.879 F1-score = 0.858	No	Individual appliance identification
Zhang et al. (2020) [14]	≤ 1 Hz	Convolutional Neural Network (CNN)	F1-score = 0.813	No	Individual appliance identification
Rafiq et al. (2020) [15]	1 Hz	Multi-Feature input Space Long Short-Term Memory (MFS-LSTM)	F1-score = 0.89 F1-score = 0.976	No	Individual appliance identification
Welikala et al. (2019) [36]	1 Hz	Feature Matching	Acc(%) = 92.50	No	Multiple appliance identification
Le et al. (2018) [16]	15 Hz	Decision Tree	F1-score = 0.92 Acc(%) = 92.64	No	Multiple appliance identification
Dinesh et al. (2015) [17]	≤ 1 Hz	Feature Matching	Acc(%) = 94.04 Acc(%) = 95.10	No	Multiple appliance identification
In this work	1 Hz	Random Forest Classification	F1-score = 0.954 Acc (%) = 94.60	Yes	Individual and Multiple appliance identification

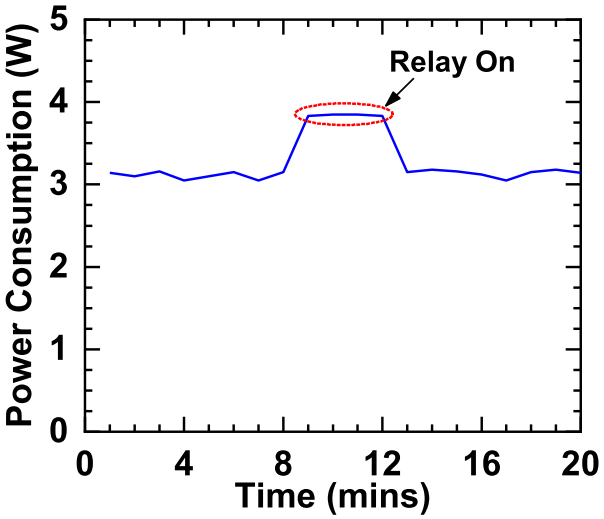


FIGURE 10. The power consumption curve of the RT-NILC system.

supports 4G, 3G, and GPRS data transmission) module can be employed.

C. COMPARATIVE STUDY

The comparison of this work with some previous NILM work at low-frequency data is shown in **TABLE 8**. Corrêa *et al.* introduced auto-associative neural network (AANN) and multi-layer perceptron (MLP) neural network algorithms on low-frequency power series signals for individual appliance identification [13]. They employed REDD and UK-DALE datasets and achieved F1-scores of 0.879 and 0.858, respectively, for the datasets. Zhang *et al.* demonstrated power series signal for individual appliance identification on REDD dataset using CNN [14]. Rafiq *et al.* reported P, Q, S, V, I, and PF as a feature for individual appliance identification using multi-feature input space long short-term memory (MFS-LSTM) algorithm [15]. They employed UK-DALE and ECO datasets and gained F1-scores of 0.89 and 0.976, respectively, for the datasets. The rms voltage and active power series data were considered to identify multiple appliances by Welikala *et al.* [36]. They developed a power matching algorithm to identify combinations of multiple appliances. An experimental dataset was prepared and had achieved an accuracy of 92.5% using a power matching algorithm. Transient features from 15 Hz power series signal were represented by Le *et al.* for multiple appliance identification using a decision tree algorithm [16]. They considered the experimental data and achieved the F1-score of 0.926 with an accuracy of 92.64%. Tracebase and REDD datasets had been employed for multiple appliance identification by Dinesh *et al.* [17]. They employed the features of power-series signal and achieved accuracy 94.04% and 95.10%, respectively, using a power matching algorithm. Though there are many works related to the NILM system at low frequency, it is difficult to compare the previously reported results because of different datasets, appliances, and devices for data preparation. This work proposed a low-cost hardware and software solution for real-time load classification (single or multiple) and electricity monitoring over IoT. The proposed system is validated in real-time using 0.2 Hz sample frequency of V, I and P data from a house. Therefore, the proposed cost-effective RT-NILM over IoT provides

a complete NILM solution, especially for low-frequency applications.

VIII. CONCLUSION

The NILM is an exceptionally cost-effective and powerful tool for load monitoring. Instead of installing multiple sensors for monitoring various electrical appliances, a single-entry point sensor instalment can yield a cost-effective and efficient solution. Therefore, the NILM idea emerges as a very promising concept for identifying individual electric appliances from a single point accumulated data. This work provides a practically feasible end-to-end hardware and software solution for non-intrusive real-time load identification over the internet. The training database building algorithm uses a novel approach to build all possible combinations of individual appliances using supply voltage-dependent signatures of each appliance. Using these training datasets, best-performed machine learning models are selected and then validated under a new dataset. The outperformed machine learning model is then incorporated with real-time load classification over IoT. After the complete implementation of RT-NILC over IoT, a final validation is performed in a real residential house. The final stage performance evaluation confirms the efficacy and feasibility of the proposed RT-NILC. Even under realistic scenarios, the proposed system has produced very accurate results with an average accuracy of more than 94%. The presented prototype system incorporates a dedicated low-cost hardware design, online database management, and software interface to facilitate user-friendly real-time operation over the internet.

ACKNOWLEDGMENT

This work was supported by Jashore University of Science and Technology (JUST), Bangladesh, under the project entitled ''Energy Saver System for Domestic Loads''. The authors, therefore, gratefully acknowledge JUST for technical and financial supports. The authors would also like to acknowledge Khulna University of Engineering and Technology (KUET), Bangladesh and Oregon Institute of Technology, USA for the collaborative support for this work.

REFERENCES

[1] G. W. Hart, ''Nonintrusive appliance load monitoring,'' *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec. 1992, doi: 10.1109/5.192069.

[2] A. Zoha, A. Gluhak, M. A. Imran, and S. Rajasegarar, ''Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey,'' *Sensors*, vol. 12, no. 12, pp. 16838–16866, Dec. 2012, doi: 10.3390/s121216838.

[3] D. Garcia-Perez, D. Perez-Lopez, I. Diaz-Blanco, A. Gonzalez-Muniz, M. Dominguez-Gonzalez, and A. A. Cuadrado Vega, ''Fully-convolutional denoising auto-encoders for NILM in large non-residential buildings,'' *IEEE Trans. Smart Grid*, vol. 12, no. 3, pp. 2722–2731, May 2021, doi: 10.1109/TSG.2020.3047712.

[4] W. Xin, B. Qi, L. Han, Z. Wang, and C. Dong, ''Fast non-intrusive load identification algorithm for resident load based on template filtering,'' *Autom. Electr. Power Syst.*, vol. 41, no. 2, pp. 135–141, 2017, doi: 10.7500/AEPS20160411001.

[5] L. Ruyi, W. Xiaohuan, H. Meixuan, Z. Hong, and H. Wenshan, ''Application of RPROP neural network in non-intrusive load decomposition,'' *Power Syst. Protection Control*, vol. 44, no. 7, pp. 55–61, 2016. [Online]. Available: http://www.epscp.com/dlbh/ch/reader/view_abstract.aspx?file_no=20160709, doi: 10.7667/PSPC151825.

[6] T.-T.-H. Le, H. Kang, and H. Kim, ''Household appliance classification using lower odd-numbered harmonics and the bagging decision tree,'' *IEEE Access*, vol. 8, pp. 55937–55952, 2020, doi: 10.1109/ACCESS.2020.2981969.

[7] Y.-H. Lin and M.-S. Tsai, ''Development of an improved time–frequency analysis-based nonintrusive load monitor for load demand identification,'' *IEEE Trans. Instrum. Meas.*, vol. 63, no. 6, pp. 1470–1483, Jun. 2014, doi: 10.1109/TIM.2013.2289700.

[8] D. Egarter, V. P. Bhuvana, and W. Elmenreich, ''PALDi: Online load disaggregation via particle filtering,'' *IEEE Trans. Instrum. Meas.*, vol. 64, no. 2, pp. 467–477, Feb. 2015, doi: 10.1109/TIM.2014.2344373.

[9] B. Liu, W. Luan, and Y. Yu, ''Dynamic time warping based non-intrusive load transient identification,'' *Appl. Energy*, vol. 195, pp. 634–645, Jun. 2017, doi: 10.1016/j.apenergy.2017.03.010.

[10] R. A. S. Fernandes, I. N. da Silva, and M. Oleskovicz, ''Load profile identification interface for consumer online monitoring purposes in smart grids,'' *IEEE Trans. Ind. Informat.*, vol. 9, no. 3, pp. 1507–1517, Aug. 2013, doi: 10.1109/TII.2012.2234469.

[11] M. Aiad and P. H. Lee, ''Non-intrusive load disaggregation with adaptive estimations of devices main power effects and two-way interactions,'' *Energy Buildings*, vol. 130, pp. 131–139, Oct. 2016, doi: 10.1016/j.enbuild.2016.08.050.

[12] A. Cominola, M. Giuliani, D. Piga, A. Castelletti, and A. E. Rizzoli, ''A hybrid signature-based iterative disaggregation algorithm for non-intrusive load monitoring,'' *Appl. Energy*, vol. 185, pp. 331–344, Jan. 2017, doi: 10.1016/j.apenergy.2016.10.040.

[13] S. D. J. do Carmo and A. R. G. Castro, ''Automated non-intrusive load monitoring system using stacked neural networks and numerical integration,'' *IEEE Access*, vol. 8, pp. 210566–210581, 2020, doi: 10.1109/ACCESS.2020.3039639.

[14] Y. Zhang, B. Yin, Y. Cong, and Z. Du, ''Multi-state household appliance identification based on convolutional neural networks and clustering,'' *Energies*, vol. 13, no. 4, p. 792, Feb. 2020, doi: 10.3390/en13040792.

[15] H. Rafiq, X. Shi, H. Zhang, H. Li, and M. K. Ochani, ''A deep recurrent neural network for non-intrusive load monitoring based on multi-feature input space and post-processing,'' *Energies*, vol. 13, no. 9, p. 2195, May 2020, doi: 10.3390/en13092195.

[16] T.-T.-H. Le and H. Kim, ''Non-intrusive load monitoring based on novel transient signal in household appliances with low sampling rate,'' *Energies*, vol. 11, no. 12, p. 3409, Dec. 2018, doi: 10.3390/en11123409.

[17] C. Dinesh, B. W. Nettasinghe, R. I. Godaliyadda, M. P. B. Ekanayake, J. Ekanayake, and J. V. Wijayakulasooriya, ''Residential appliance identification based on spectral information of low frequency smart meter measurements,'' *IEEE Trans. Smart Grid*, vol. 7, no. 6, pp. 2781–2792, Nov. 2016, doi: 10.1109/TSG.2015.2484258.

[18] C. C. Yang, C. S. Soh, and V. V. Yap, ''A non-intrusive appliance load monitoring for efficient energy consumption based on naive Bayes classifier,'' *Sustain. Comput., Informat. Syst.*, vol. 14, pp. 34–42, Jun. 2017, doi: 10.1016/j.suscom.2017.03.001.

[19] M. Figueiredo, B. Ribeiro, and A. de Almeida, ''Electrical signal source separation via nonnegative tensor factorization using on site measurements in a smart home,'' *IEEE Trans. Instrum. Meas.*, vol. 63, no. 2, pp. 364–373, Feb. 2014, doi: 10.1109/TIM.2013.2278596.

[20] O. Parson, S. Ghosh, M. Weal, and A. Rogers, ''An unsupervised training method for non-intrusive appliance load monitoring,'' *Artif. Intell.*, vol. 217, pp. 1–19, Dec. 2014, doi: 10.1016/j.artint.2014.07.010.

[21] H.-H. Chang, P. W. Wiratha, and N. Chen, ''A non-intrusive load monitoring system using an embedded system for applications to unbalanced residential distribution systems,'' *Energy Procedia*, vol. 61, pp. 146–150, Jan. 2014, doi: 10.1016/j.egypro.2014.11.926.

[22] J. Kelly and W. Knottenbelt, ''Neural NILM: Deep neural networks applied to energy disaggregation,'' in *Proc. 2nd ACM Int. Conf. Embedded Syst. Energy-Efficient Built Environ.*, New York, NY, USA, Nov. 2015, pp. 55–64, doi: 10.1145/2821650.2821672.

[23] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajic, ''AMPds: A public dataset for load disaggregation and eco-feedback research,'' in *Proc. IEEE Electr. Power Energy Conf.*, Halifax, NS, Canada, Aug. 2013, pp. 1–6, doi: 10.1109/EPEC.2013.6802949.

[24] P. Schirmer and I. Mporas, ''Statistical and electrical features evaluation for electrical appliances energy disaggregation,'' *Sustainability*, vol. 11, no. 11, p. 3222, Jun. 2019, doi: 10.3390/su11113222.

[25] J. He, Z. Zhang, L. Zhu, Z. Zhu, J. Liu, and K. Gai, "An efficient and accurate nonintrusive load monitoring scheme for power consumption," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 9054–9063, Oct. 2019, doi: 10.1109/JIOT.2019.2926815.
[26] M. M. Hasan, D. Chowdhury, and M. Z. R. Khan, "Non-intrusive load monitoring using current shapelets," *Appl. Sci.*, vol. 9, no. 24, pp. 53–63, Dec. 2019, doi: 10.3390/app9245363.
[27] P. Meehan, C. McArdle, and S. Daniels, "An efficient, scalable time-frequency method for tracking energy usage of domestic appliances using a two-step classification algorithm," *Energies*, vol. 7, no. 11, pp. 7041–7066, Oct. 2014, doi: 10.3390/en7117041.
[28] H.-H. Chang, K.-L. Lian, Y.-C. Su, and W.-J. Lee, "Power-spectrum-based wavelet transform for nonintrusive demand monitoring and load identification," *IEEE Trans. Ind. Appl.*, vol. 50, no. 3, pp. 2081–2089, May/Jun. 2014, doi: 10.1109/TIA.2013.2283318.
[29] L. Guo, S. Wang, H. Chen, and Q. Shi, "A load identification method based on active deep learning and discrete wavelet transform," *IEEE Access*, vol. 8, pp. 113932–113942, 2020, doi: 10.1109/ACCESS.2020.3003778.
[30] Z. Zheng, H. Chen, and X. Luo, "A supervised event-based non-intrusive load monitoring for non-linear appliances," *Sustainability*, vol. 10, no. 4, p. 1001, Mar. 2018, doi: 10.3390/su10041001.
[31] H.-H. Chang, "Non-intrusive demand monitoring and load identification for energy management systems based on transient feature analyses," *Energies*, vol. 5, no. 11, pp. 4569–4589, Nov. 2012, doi: 10.3390/en5114569.
[32] Q. Wu and F. Wang, "Concatenate convolutional neural networks for non-intrusive load monitoring across complex background," *Energies*, vol. 12, no. 8, p. 1572, Apr. 2019, doi: 10.3390/en12081572.
[33] F. Ciancetta, G. Bucci, E. Fiorucci, S. Mari, and A. Fioravanti, "A new convolutional neural network-based system for NILM applications," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, 2021, Art. no. 1501112, doi: 10.1109/TIM.2020.3035193.
[34] J. Kim, T.-T.-H. Le, and H. Kim, "Nonintrusive load monitoring based on advanced deep learning and novel signature," *Comput. Intell. Neurosci.*, vol. 2017, pp. 1–22, Oct. 2017, doi: 10.1155/2017/4216281.
[35] H.-H. Chang, L.-S. Lin, N. Chen, and W.-J. Lee, "Particle-swarm-optimization-based nonintrusive demand monitoring and load identification in smart meters," *IEEE Trans. Ind. Appl.*, vol. 49, no. 5, pp. 2229–2236, Sep./Oct.2013, doi: 10.1109/TIA.2013.2258875.
[36] S. Welikala, N. Thelasingha, M. Akram, P. B. Ekanayake, R. I. Godaliyadda, and J. B. Ekanayake, "Implementation of a robust real-time non-intrusive load monitoring solution," *Appl. Energy*, vol. 238, pp. 1519–1529, Mar. 2019, doi: 10.1016/j.apenergy.2019.01.167.
[37] X. Sun, X. Wang, Y. Liu, and J. Wu, "Non-intrusive sensing based multi-model collaborative load identification in cyber-physical energy systems," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (IMTC)*, Montevideo, Uruguay, May 2014, pp. 1–6, doi: 10.1109/I2MTC.2014.7461656.
[38] F. C. C. Garcia, C. M. C. Creayla, and E. Q. B. Macabebe, "Development of an intelligent system for smart home energy disaggregation using stacked denoising autoencoders," *Procedia Comput. Sci.*, vol. 105, pp. 248–255, Jan. 2017, doi: 10.1016/j.procs.2017.01.218.
[39] X.-C. Le, B. Vrigneau, and O. Sentieys, "L1-norm minimization based algorithm for non-intrusive load monitoring," in *Proc. IEEE Int. Conf. Pervas. Comput. Commun. Workshops (PerCom Workshops)*, St. Louis, MO, USA, Mar. 2015, pp. 299–304, doi: 10.1109/PERCOMW.2015.7134052.
[40] Y.-Y. Chen, Y.-H. Lin, C.-C. Kung, M-H. Chung, and I. Yen, "Design and implementation of cloud analytics-assisted smart power meters considering advanced artificial intelligence as edge analytics in demand-side management for smart Homes," *Sensors*, vol. 19, no. 9, p. 2047, May 2019, doi: 10.3390/s19092047.
[41] P. Franco, J. M. Martinez, Y.-C. Kim, and M. A. Ahmed, "IoT based approach for load monitoring and activity recognition in smart Homes," *IEEE Access*, vol. 9, pp. 45325–45339, 2021, doi: 10.1109/ACCESS.2021.3067029.
[42] *PZEM-004t, Datasheet*. Accessed: May 15, 2021. [Online]. Available: https://innovatorsguru.com/wp-content/uploads/2018/01/PZEM-004T-Datasheet.pdf
[43] *Fluke-5502A, Specification*. Accessed: May 15, 2021. [Online]. Available: http://download.flukecal.com/pub/literature/4225366C_w.pdf
[44] *Owl USB CM160 Electricity Monitor*. Accessed: May 15, 2021. [Online]. Available: https://www.amazon.co.uk/dp/B004BDNR84
[45] *Metered Variac—Variable AC Output Transformer, Specification*. Accessed: May 15, 2021. [Online]. Available: https://satkit.com/en-gb/variable-output-transformer-variac-ac-12-amp-0-250v-tdg
[46] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, and J. Vanderplas, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.
[47] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

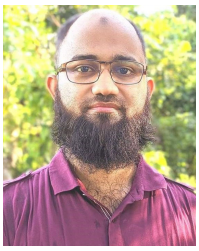
MD. TANVIR AHAMMED received the B.Sc. and M.Sc. degree in electrical and electronic engineering (EEE) from Khulna University of Engineering and Technology (KUET), Bangladesh, in 2015 and 2019, respectively. He is currently working as an Assistant Professor with the Department of Electrical and Electronic Engineering, Jashore University of Science and Technology (JUST), Bangladesh. He has published his research papers in national and international conferences. His research interests include machine learning, deep learning, the IoT, and smart grid.

MD. MEHEDI HASAN received the B.Sc. degree in electrical and electronic engineering (EEE) from Bangladesh University of Engineering and Technology (BUET), Bangladesh, in 2016. He is currently working as an Assistant Professor with the Department of Electrical and Electronic Engineering, Jashore University of Science and Technology (JUST), Bangladesh. He has published his research papers in national and international conferences and journals. His research interests include machine learning, deep learning, and computer vision.

MD. SHAMSUL AREFIN (Member, IEEE) received the B.Sc. degree in electronics and telecommunication engineering (ETE) from Daffodil International University (DIU), Bangladesh, in 2009, and the Ph.D. degree in engineering from Swinburne University of Technology (SUT), Australia, in 2015. He is currently working as an Assistant Professor with the Department of Electrical and Electronic Engineering (EEE), Bangladesh University of Business and Technology (BUBT), Bangladesh. He has been working in multidisciplinary engineering, such as the IoT, AI, distributed power systems, telecommunication, and biomedical for last six years and published several research papers including couple of book chapters. He is also working as an External Consultant with the Control and Applications Research Centre (CARC), BRAC University, Bangladesh, and successfully completed multiple projects, in the field of the IoT, renewable energy, and biomedical applications.

MD. RAFIQUL ISLAM (Senior Member, IEEE) received the B.Sc. degree in electrical and electronic engineering from Bangladesh University of Engineering and Technology (BUET), in 1998, the M.Sc. degree from Khulna University of Engineering and Technology (KUET), Bangladesh, in 2006, and the Ph.D. degree in semiconductor device growth, characterization, and fabrication from the University of Fukui, Japan, in 2010.

He is currently working as a Professor with Khulna University of Engineering and Technology, Khulna, Bangladesh. He received the Monbukagakusho Scholarship from MEXT, Japan, from 2007 to 2010. He has published around 100 research papers in national and international conferences and journals. His research interests include thin films solar cells, growth characterization and fabrications, advanced semiconductor materials properties, and compound semiconductor-based devices. His two papers received the best paper award at international conferences.



MD. AMINUR RAHMAN received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from Khulna University of Engineering and Technology (KUET), Bangladesh, in 2009 and 2014, respectively, where he is currently pursuing the Ph.D. degree, as a part-time student. He is currently working as the Deputy Director (Project) with the Bangladesh Power Management Institute (BPMI) of Power Division, Ministry of Power, Energy and Mineral Resources, Bangladesh, where he is involved as a coordinator with several technical trainings, projects, and research works at the Power Sector of Bangladesh. He is the author or coauthor of eight research papers in conferences and journals. His research interests include strain impacts in solar cell, renewable energy integration, and smart grid and power system stability.

EKLAS HOSSAIN (Senior Member, IEEE) received the B.S. degree in electrical and electronic engineering from Khulna University of Engineering and Technology, Bangladesh, in 2006, the M.S. degree in mechatronics and robotics engineering from the International Islamic University of Malaysia, Malaysia, in 2010, and the Ph.D. degree from the College of Engineering and Applied Science, University of Wisconsin-Milwaukee (UWM). He has been working in the area of distributed power systems and renewable energy integration for last ten years and he has published a number of research papers and posters in this field. He has been involved in several research projects on renewable energy and grid tied microgrid system at Oregon Tech, as an Assistant Professor with the Department of Electrical Engineering and Renewable Energy, since 2015. He is a Senior Member of the Association of Energy Engineers (AEE). He is currently serving as an Associate Editor for IEEE Access. He is working as an Associate Researcher at Oregon Renewable Energy Center (OREC). He is a registered Professional Engineer (PE) in the state of Oregon, USA. He is also a Certified Energy Manager (CEM) and a Renewable Energy Professional (REP). His research interests include modeling, analysis, design, and control of power electronic devices; energy storage systems; renewable energy sources; integration of distributed generation systems; microgrid and smart grid applications; robotics; and advanced control systems. He is the winner of the Rising Faculty Scholar Award from Oregon Institute of Technology for his outstanding contribution in teaching, in 2019. He with his dedicated research team is looking forward to exploring methods to make the electric power systems more sustainable, cost-effective, and secure through extensive research and analysis on energy storage, microgrid systems, and renewable energy sources.

MD. TANVIR HASAN (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from Khulna University of Engineering and Technology (KUET), Bangladesh, in 2006 and 2007, respectively, and the Ph.D. degree in electrical and electronic engineering from the Graduate School of Engineering, University of Fukui, Japan, in 2013.

He is currently working as an Associate Professor with the Department of Electrical and Electronic Engineering, Jashore University of Science and Technology (JUST), Bangladesh. He received the Monbukagakusho Scholarship from MEXT, Japan, from 2010 to 2013. He served as a Secretary, a IEEE Young Professional, Bangladesh Section (BDS), and an Executive Committee Member and as a Professional Activity Coordinator at IEEE BDS, in 2016 and 2017, respectively. He is the author or coauthor of more than 50 research papers in conferences and journals. His research interests include growth, design, fabrication, characterization, simulation, and modeling of III-V-based semiconductor devices (electronic and optoelectronic). He received the IEEE Student Paper Award (Honorary Mentioned) from the IEEE Electron Devices Society, Bangladesh Chapter, in 2007. He has been a member of the IEEE Electron Devices Society, since 2008. He serves as a Reviewer for IEEE Transactions on Electron Devices, *Journal of Applied Physics*, and *Applied Physics Letters*.