Machine Learning-Based Malicious User Detection in Energy Harvested Cognitive Radio-Internet of Things

Md. Sipon Miah, Mohammad Amzad Hossain, Kazi Mowdud Ahmed, Md. Mahbubur Rahman, Ali Calhan, and Murtaza Cicioglu

***Abstract*—The Internet of things (IoT) is a network of interconnected objects that are connected and controls autonomous machines in the world. The cognitive radio based Internet of things (CR-IoT) concept is a revolutionary technology for the future of IoT that mitigates the spectrum scarcity problem. However, each CR-IoT user does not obtain a better sensing gain, an enhanced sum rate, and a prolonged network lifetime in conventional CR-IoT networks under the existing energy harvesters due to both (i) underutilizing the reporting framework and (ii) without separating normal and abnormal (malicious) CR-IoT users. For these reasons, we proposed machine learning (e.g. logistic regression (LR), support vector machine (SVM) and k-nearest neighbors (k-NN)) based malicious user detection in energy harvested CR-IoT networks, where each CR-IoT user will be powered by finite capacity batteries and energy harvesters. The main contributions of this paper: First, we reviewed the technological attributes and platforms proposed in the current literature for the sensing, sum rate, and network lifetime with security threats; Second, the proposed classification algorithms using machine learning are divided into two groups between normal and abnormal (malicious) CR-IoT users; Third, this scheme is utilized the reporting framework by only normal CR-IoT users where each normal CR-IoT user is obtained a longer sensing time slot; Fourth, as a proof-of-concept, the performance of the proposed scheme is evaluated through numerical experiment; Finally, this proposed scheme is greatly achieved better sensing gain, enhanced sum rate, and prolonged network lifetime, in comparison to the existing conventional schemes.**

M. S. Miah is with the Department of Information and Communication Technology, Islamic University, Kushtia-7003, Bangladesh, and also with the Department of Computer Science, National University of Ireland Galway, Galway, Republic of Ireland e-mail: m.miah1@nuigalway.ie (www.nuigalway.ie).

M. A. Hossain is with the Department of Computer Science, National University of Ireland Galway, Galway, Republic of Ireland e-mail: m.hossain3@nuigalway.ie (www.nuigalway.ie).

K. M. Ahmed is with the Department of Information and Communication Technology, Islamic University, Kushtia-7003, Bangladesh e-mail: mowdud@ice.iu.ac.bd(www.iu.ac.bd).

M. M. Rahman is with the Department of Information and Communication Technology, Islamic University, Kushtia-7003, Bangladesh e-mail: mrahman@ice.iu.ac.bd(www.iu.ac.bd).

A. Calhan is with the Department of Computer Engineering, Duzce University, Duzce, Turkey e-mail: alicalhan@duzce.edu.tr(www.duzce.edu.tr).

M. Cicioglu is with Information Technologies Department, Ministry of National Education, Bolu, Turkey e-mail: murtazacicioglu@gmail.com(www.bolu.meb.gov.tr).

Manuscript received January 19, 2021; revised August 26, 2021.

***Index Terms*—Cognitive radio, internet of things, machine learning, spectrum sensing, detection performance, sum rate, network lifetime.**

I. Introduction

INTERNET of things (IoT) consists of objects or things that are meant to sense their surrounding environments, exchange, and communicate information or data with others devices (e.g., computers, mobiles, cordless phones, any wired or wireless devices) with different applications [1], [2]. The number of devices connected to either public or private networks via wired or wireless has recently increased significantly [3]. However, this dramatic increase in IoT causes spectrum shortage that will be resulting in degradation in detection performance because of the massive number of IoT devices are connecting to the licensed band.

Since the last decade, the idea of cognitive radio (CR) for the design of wireless communications networks has been developed to alleviate the shortage problem of insufficient radio spectrum by enhancing spectrum utilization [4]–[6]. The integration of CR and IoT i.e., CR based IoT (CR-IoT) is in the developing field of future wireless communications which have the great potential to opportunistically access the allocated spectrum bands (licensed) and support new IoT services (e.g., traffic related to disaster management, banking, response planning, security, healthcare, agriculture, and education) that can profoundly impact our lives in a positive way. Therefore, CR-IoT networks are becoming an attractive solution for spectrum scarcity problem, low sum rate, and a shorter network lifetime [7]–[14]. Furthermore, energy harvesting is becoming increasingly important to complement existing battery-powered wireless communication networks, extending their longevity, and keeping them more environmental-friendly. In addition, it can prolong network lifetime by applying energy harvesting techniques to a machine learning-based malicious user detection in energy harvesting CR-IoT networks [15]. Despite the potential benefits of CR-IoT networks, the major challenges faced by network researchers and engineers of IoT networks are to better sensing gain, enhance sum rate, and prolong network lifetime of a typical CR-IoT users or secondary users (SUs) (e.g., consisting of malicious users) with accommodating increasingly growing new applications and services more

than a limited available spectrum band. For example, the sensing gain of the CR-IoT network drops significantly when multiple malicious users are sensing the licensed spectrum band. Sensing gain and classification algorithm of these networks are required for efficient detection and deployment of such systems.

The machine learning (ML) based malicious user detection in energy harvested CR-IoT networks has recently attracted significant interest in the wireless research community [16]–[22]. The authors are concerned with separating malicious users in energy harvested CR-IoT networks, as specified in the ML model [23]. Specifically to be addressed is the potential of the machine learning to make a CR-IoT network better to identify malicious CR-IoT users. The main contributions of this paper are

- We proposed a novel classification algorithm based on the machine learning, i.e., logistic regression (LR), support vector machine (SVM), and k-nearest neighbors (k-NN) for separating normal CR-IoT users and malicious CR-IoT users.
- The sensing performance is performed by the energy detection technique, the proposed ML algorithms are employed on the data set and then it makes group for normal CR-IoT users and abnormal CR-IoT users (malicious users).
- After the grouping normal CR-IoT users and malicious CR-IoT users at the fusion centre (FC), the FC needs to employ the Dempster-Shafer (DS) theory to evaluate the sensing performance of the proposed ML algorithms.
- The sum rate of a PU network and the CR-IoT network, including only normal CR-IoT users, is evaluated on the basis of the sensing performance of the proposed ML algorithms.
- Moreover, for the proposed ML algorithms, the prolonged network life of the energy harvested CR-IoT network, including only normal CR-IoT users, is evaluated based on the sensing performance and sum rate.
- The results of the simulation demonstrate that the proposed scheme is achieved an improved sensing performance, a better sum rate, and a prolonged network lifetime of the energy harvested CR-IoT networks compared to other conventional schemes.

The remainder of this paper is structured as follows. In Section II, the related works with contributions are addressed. The proposed system model with explanation is presented in Section III. Analysis of different network metrics based on machine learning algorithms e.g., LR, SVM and k-NN is discussed in Section IV. The results of the simulation of the proposed scheme is validated in Section V. In Section VI, closing notes for possible works are point-outs.

II. Related Works

The CR is in the developing field of cognitive radio network (CRN) [6] which have great potential to allow CR users to access the allocated spectrum bands which are temporally idle due to each CR user being equipped with energy harvesting. The IoT is a modern machine-to-machine (M2M) communication paradigm that allows machines, computers, devices, apps, or appliances without human interference to communicate with each other [7]. In CR-IoT networks, each CR-IoT device/user shares information of different objects like the things-oriented, Internet-oriented and semantic-oriented over the Internet using different communication technologies [9]–[11], [13], [14]. In the current era, the dramatic growth of the number of CR-IoT users that are needed more available spectrum. However, the scarcity of the available spectrum in a CR-IoT network is a major challenge due to the spectrum allocation techniques in a CR-IoT network which are totally under utilized by CR-IoT users. Therefore, for CR-IoT networks, efficient spectrum sensing is essential. The cooperative spectrum sensing (CSS) approach is that CR-IoT users share the sensing results with the fusion center (FC) to make decisions. With this approach, hidden PU offers a higher and more reliable sensing opportunity for the problem compared to individual sensing performance. The CSS approach is sensitive to attacks by malicious users who transmit incorrect sensing results to the FC [15]. However, the sum rate and the network lifetime are not analyzed in energy harvested CR-IoT networks. Robust machine learning (ML) based spectrum sensing in cognitive radio networks (CRNs) proposed by Shah et al. [16], where each CR user contrasts their current sensing data to existing sensing class and measures distance vectors. However, the malicious user detection, sum rate, and network lifetime were not analyzed. Jan et al. [17] proposed a spectrum sensing architecture with multi-class hypotheses in CRNs, where each CR user enhances throughput using an support vector machine (SVM). However, the malicious user detection and network lifetime were not analyzed. Zhu et al. [18] proposed a new Q-learning based transmission scheduling framework using deep learning in a CR-IoT network, where each CR-IoT user is maximizing the system throughput using the required strategy to transfer packets in different buffers over multiple channels. However, the malicious user detection and network lifetime were not analyzed. Rahman et al. [19] proposed reinforcement learning based on efficient transmission mode selection for cooperative CRNs, where the proposed scheme is to analyze the energy efficiency, time delay, and PU interference. However, the malicious user detection, sum rate, and network lifetime were not analyzed. Thilina et al. [20] proposed a ML technique for CSS in CRNs, where the proposed scheme is analysis of the detection performance based on unsupervised ML and supervised ML. However, the malicious user detection, sum rate, and network lifetime were not analyzed. Mustafa et al. [21] proposed a survey of ML algorithms and their CR applications, where many ML algorithms and their CR applications are discussed in the proposed scheme with regards of the detection performance, classification of modulation, and allocation of power. Hung et al. [22] proposed a fuzzy SVM algorithm for CSS with noise

uncertainty, where the parametric tests obtained by the secondary users are arranged into a feature vector instead of being combined through weighted sum. However, the malicious user detection, sum rate, and network lifetime were not analyzed. Li et al. [23] proposed an improved CSS model based on ML for CRNs that utilizes user classification methods to minimize overhead cooperation and efficiently enhance detection performance. However, the sum rate and the network lifetime were not analyzed. Moreover, the harvested energy from the CR-IoT users were not analyzed. In [24], the authors proposed a ML framework for detecting the PU emulation attack (PUEA) in CRNs, where various classification algorithms used to differentiate between malicious users (MUs) and legitimate users (LUs). However, the sensing performance, sum rate, and network lifetime were not evaluated. In [25], the authors proposed ML approach for CRNs, where unsupervised ML, supervised ML, semi-supervised ML, deep learning algorithms used to detect the PUEA and improve detection performance. However, the sum rate and the network lifetime were not analyzed. In [26], the authors proposed SVM algorithm based malicious primary user emulation signal detection model, which classifies the PU and the malicious PU signal while using the signal-to-noise ratio (SNR) and energy signal entropy. This model is enhanced the detection performance of the existence of MUs in low SNR even without a threshold calculation and detection probability of the legitimate PU. However, the sum rate and the network lifetime were not analyzed. In summary, the current research has some drawbacks as shown in Table I: (i) a typical CR-IoT network in which all CR-IoT users including normal and malicious CR-IoT users are participating to sense the PU licensed channel which not to enhance the sensing gain; (ii) an enhanced the sum rate and a prolonged network lifetime has not been analyzed regarding the detection probability, the probability of false alarm, and the energy harvested from the PU licensed channel. The proposed scheme overcome these drawbacks.

III. Proposed System Model

In this system model, we proposed the ML algorithms based malicious CR-IoT users detection in energy harvested CR-IoT networks as shown in Fig. 1 where to separate the normal CR-IoT users and malicious CR-IoT users of a CR-IoT network operating in both with fading and without fading phenomena. As in Fig. 1, we consider CR-IoT networks with N normal CR-IoT users and M malicious CR-IoT users. In the training phase as shown in Fig. 1, the operating environment is discovered by measuring the action of the CR-IoT user with the changing of the PU activities. Each CR-IoT user sends their signal to noise ratio (SNR) to the FC and then the FC separates the normal and malicious CR-IoT users based on the ML algorithms i.e., LR, SVM, and k-NN during the classification phase where all malicious CR-IoT users are dropout due to this malicious CR-IoT users can severely reduce the sensing gain. In sensing phase, only each normal

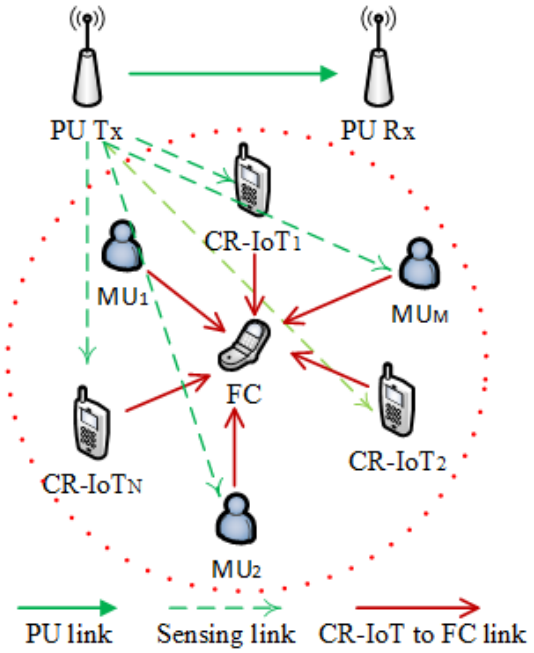

Fig. 1. The proposed system model here the normal CR-IoT users (N) and the malicious CR-IoT users (M).

CR-IoT user (e.g., i^{th} CR-IoT user) generates a sensing report during the longer sensing time (due to utilizing the reporting framework) which makes a local decision, and transmits this local decision during the fixed reporting time to the FC which will be combined to make a global decision about the PU activities on the licensed channel. In the secondary CR-IoT network, there is a control channel that is used to communicate data from CR-IoT users (e.g., normal and abnormal CR-IoT users) to FC. It is assumed that the reporting signal is fading or error-free. In data transmission phase, the source CR-IoT user (e.g. CR-IoT $\mathbb{T}$x) uses a direct connection to transmit the data to the destination CR-IoT user (CR-IoT $\mathbb{R}$x) during the rest of time slot. Various factors are analyzed, including the number of CR-IoT users (e.g., normal users and malicious users) and their traffic load distribution, the changing activities of the PU, and wireless channel conditions (e.g., without fading and with fading) that affect the detection performance of an energy harvested CR-IoT network.

A. Primary Users Traffic Model

The PU traffic is modeled as a basic mechanism of ON and OFF involving two instantaneous states e.g. $s_t = 1$ is for active and $s_t = 0$ is for inactive at the time slot t [2].

B. Cooperative Energy Vectors

Each CR-IoT user executes the local spectrum sensing independently in the PU and CR-IoT user link to identify the activities of PU signal and transmits its local report based on a binary hypothesis that is either $\mathbb{H}_0$ or $\mathbb{H}_1$; here, $\mathbb{H}_0$ and $\mathbb{H}_1$ are the absence of PU and the presence of PU,

TABLE I
COMPARISON OF EXISTING ML ALGORITHMS TO PROPOSED ML ALGORITHMS BASED CR-IOT NETWORKS

Approaches	Proposed idea	ML	Sensing performance	Sum rate	Expected lifetime
A. Mustafa et al. (2015) [21]	A survey of ML algorithms and their applications in CR	SVM	✓	✗	✗
Y. D. Huang et al. (2017) [22]	A fuzzy SVM algorithm for CSS with noise uncertainty	SVM	✓	✗	✗
Z. Li et al. (2018) [23]	An enhanced CSS model based on ML for CRNs	SVM	✓	✗	✗
A. Albehadili et al. (2019) [24]	ML based PU emulation attack detection in CRNs using pattern described link-signature (PDLS)	ML	✓	✗	✗
A. Albehadili et al. (2019) [25]	Semi-supervised ML for PU emulation attack detection and prevention through core based analytic for CRNs	ML	✓	✗	✗
E. C. Munoz et al. (2020) [26]	Detection of Malicious PU emulation Based on a SVM for a mobile CRN using software defined radio	SVM	✓	✗	✗
M. S. Khan et al. (2020) [27]	SVM based classification of MUs in CRNs	SVM	✓	✗	✗
M. S. Khan et al. (2020) [28]	A genetic algorithm based soft decision fusion scheme in cognitive IoT networks with MUs	GA	✓	✗	✗
Y. Zhang et al. (2020) [29]	Ensemble learning based robust CSS in full-duplex CRNs	SVM, LR	✓	✗	✗
M. S. Miah et al. (proposed)	ML based MU detection in energy harvested CR-IoT	SVM, k-NN, LR	✓	✓	✓

respectively. The spectrum sensing problem of the CR-IoT user can be computed based on a binary hypothesis as follows:

$$\begin{cases} \mathbb{H}_0: & \text{if PU is for inactive at time slot } (s_t = 0); \\ \mathbb{H}_1: & \text{if PU is for active at time slot } (s_t = 1); \end{cases} \tag{1}$$

The signals received of the i^{th} CR-IoT user can be computed according to the packet transmission of the PU [30] as follows:

$$z_i(t) = \begin{cases} y_i(t); & \mathbb{H}_0 \\ h_i(t)\, x(t) + y_i(t); & \mathbb{H}_1 \end{cases} \tag{2}$$

where $t = 1, 2, 3, ..., L$, here L denotes the number of samples of the signals received that defines as $L = 2\tau_s f_s$; τ_s denotes the flexible sensing time slot in sec, and f_s denotes the sampling frequency. We can consider that the flexible sensing time,τ_s is proportional to the energy harvested, e_s with a constant, ζ i.e., $\tau_s = \zeta e_s$. Moreover, $z_i(t)$ denotes the signal received by the i^{th} CR-IoT user, $x(t)$ denotes the PU transmitted signal, i.e., $x(t) \sim \aleph\left(0, \sigma_x^2\right)$, and $y_i(t)$ denotes the additive white Gaussian noise of the i^{th} CR-IoT user, i.e., $y_i(t) \sim \aleph(0, \sigma_{y,i}^2)$. Also, $h_i(t)$ denotes the channel gain of the i^{th} CR-IoT user, $\mathbb{H}_1$ denotes the active of PU signal, and $\mathbb{H}_0$ denotes the inactive of PU signal.

For the CSS of the CR-IoT users, the Energy Detector (ED) technique is commonly used because it can be applied effectively without any previous PU signal information being acquired. The sensing result $z_i(t)$ obtained by the i^{th} CR-IoT user transmitter is the signal power in the time domain at a given frequency; a band-pass filter is added to the received signal, then an analog-to-digital converter (ADC) converts the output of this filter, which is independently averaged and squared using the conventional ED technique to determine its own calculated energy, E_i as follows [2]:

$$E_i = \frac{1}{L} \sum_{t=1}^{L} \| z_i(t) \|^2 \tag{3}$$

Based on the central limit theorem (CLT), when L is relatively high, the signal received at the i^{th} CR-IoT user, E_i in (3) can be estimated as a Gaussian random variable under a binary hypotheses i.e., $\mathbb{H}_0$ and $\mathbb{H}_1$ with mean and variance, respectively as follows [30]:

$$E_i \sim \begin{cases} \aleph\left(\mu_i(\mathbb{H}_0), \sigma_i^2(\mathbb{H}_0)\right) \\ \aleph\left(\mu_i(\mathbb{H}_1), \sigma_i^2(\mathbb{H}_1)\right) \end{cases} \tag{4}$$

where

$$\begin{aligned} \mu_i(\mathbb{H}_0) &= 2\tau_s f_s \sigma_{y,i}^2, \\ \sigma_i^2(\mathbb{H}_0) &= 4\tau_s f_s \sigma_{y,i}^4, \\ \mu_i(\mathbb{H}_1) &= 2\tau_s f_s \left(1 + |h_i|^2 \gamma_i\right) \sigma_{y,i}^2, \\ \sigma_i^2(\mathbb{H}_1) &= 4\tau_s f_s \left(1 + 2|h_i|^2 \gamma_i\right) \sigma_{y,i}^4, \end{aligned}$$

where γ_i denotes a signal to noise ratio (SNR) which defines as $\gamma_i = \frac{p_x^2}{\sigma_{y,i}^2}$; here p_x^2 is the signal power of $x(t)$.

IV. Analysis of different network metrics based on ML algorithms

ML is a sub-branch of computer science that was developed in 1959 from computational learning experiments and model recognition in artificial intelligence. ML is a framework that can learn as a structural function and investigate the work and construction of algorithms that can make predictions about data. ML algorithms operate by constructing a model from sample inputs to make

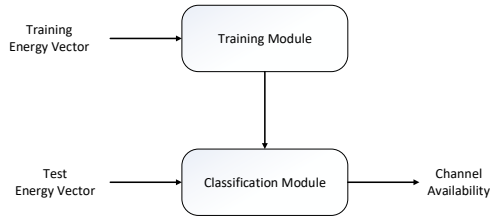

Fig. 2. Modular framework of the proposed CSS model [23].

data-based predictions and decisions, rather than simply following classic instruction set. There are three types of learning strategies in ML. ML algorithms are often used in model classification. The relevant feature vector is extracted from a model and given to the classifier so that it can be categorized. In this context, we considered an estimated energy level (energy vector) of the CR device as a feature vector. Next, the classifier transfers the energy vector to one of two classes: "CR-IoT/SU user" and "MU". First, the classifier must go through a training phase where it learns from the training of feature vectors i.e., energy vector as shown in Fig. 2.

In supervised ML, data is collected from interactive systems and arranged in a certain order. Unsupervised ML is another ML method with no data set and no outputs. It is to investigate common points by interpreting the data in the data set and to obtain meaningful data by clustering them. Semi-supervised ML is both learning that is supervised ML and unsupervised ML. It consists of using too much untagged data and small-sized data tagged together. In this paper, we use supervised ML algorithms such as LR, SVM, and k-NN for classification of a CR-IoT/SU users is either normal CR-IoT users or malicious users. The dataset for ML algorithms are constructed from sensing energies of CR-IoT/SU users and MUs. Owing to energy detection based spectrum sensing the related energy levels of users can be detected. The range of sensing energies at the valid CR-IoT/SU users is between 90 and 108 and the values outside this range can belong to invalid CR-IoT users as MUs. Benefits of the proposed solution as follows:

- Thanks to the optimized classifier trained with energy vectors, a structure capable of more comfortable adaptation to dynamic radio environments has been created. Since the training process does not require any prior knowledge and parameter setting about the environment, an environment-independent model was created.
- The proposed MU detection technique also performs better in terms of false alarm and detection probabilities, as it can make optimized decisions compared to traditional approaches.

A. Logistic Regression (LR) Algorithm

The LR algorithm in ML framework is a widely used efficient meta learner that automatically learns the optimal weights of the results of basic learner prediction [29]. LR is a regression method which is used for classification problems based on the concept of probability. It is used for the classification categorical or numerical data. It operates only if the dependent variable, i.e. the result can be taken two different values. In LR algorithm, the logistic regression function is defined as follows:

$$y = logit(p) = ln\left(\frac{p}{p-1}\right) = w^T x + b \tag{5}$$

where logistic ($logit$) is the ratio of class probabilities, x is the data vector, $y \in \{-1, +1\}$ for two classes, w and b are the weight parameters.

After training phase with the aim of finding parameters for class probabilities i.e., $Class1$ and $Class2$ are defined as follows:

$$p(y = -1|x) = 1 - p(y = +1|x) \tag{6}$$

and

$$p(y = +1|x) = 1 - p(y = -1|x) \tag{7}$$

Similarly, for class probability of C is defined as follows:

$$p(y = C|x) = \frac{1}{1 + e^{(-y(w^T x + b))}} \tag{8}$$

The benefits of using LR include its flexibility, reliability and the ability to resist over-fitting without any hyper-parameter tuning in small-scaled data sets. The LR based MUs classification ML algorithm in an energy harvested CR-IoT network is shown in Algorithm 1 which consists of four stages: data generation, sensing, classification, and area under curve (AUC).

B. Support Vector Machine (SVM) Algorithm

The SVM can be defined as a ML approach based on vector space that finds a boundary decision between two classes that are furthest from any point in the training dataset. It is a supervised ML algorithms which can be used for classification. In classification problem, it creates a set of hyper-planes in a high dimensional feature space by analyzing data for binary classification [31]. In the CR-IoT network, we consider CR-IoT networks with N normal CR-IoT users, M malicious CR-IoT users and total number of user is represent by $K = N + M$.

The notation of training data set ($\mathbb{X}$) of the proposed SVM algorithm is defined as follows [27]:

$$\mathbb{X} = \left[(x_i, y_i)\,|x_i \in \mathbb{R}^Z, y_i \in [+1, -1]\right]_{i=1}^{Z} \tag{9}$$

Here, (x_i, y_i) represents the data set for normal and malicious users which is defined as follows

$$(x_i, y_i) = ((x_1, y_1), (x_2, y_2), ..., (x_Z, y_Z)) \tag{10}$$

where x_i represents the energy vector/training data of $Z(i = 1, 2, ..., Z)$ users, $y_i \in [+1, -1]$ is the class vector/target output, and class "+1" and "-1" represent normal CR-IoT users and malicious CR-IoT users, respectively.

Algorithm 1 LR based MUs classification ML algorithm in an energy harvested CR-IoT network.

Input: θ, L and ϵ
Output: Normal CR-IoT users and Malicious Users
Initialisation :
Sensing the data
for $i = 1$ to N **do**
  for $t = l$ to L **do**
    Energy reported by the i^{th} CR-IoT users based on Eq. 3
  end for
end for
for $i = 1$ to M **do**
  for $t = l$ to L **do**
    Energy reported by the i^{th} malicious users based on Eq. 3
  end for
end for
Data processing for LR:
Combining the both users data in to one data set
Find the number of features from data set
Extract the features matrix X and label vector Y
Scaling the features data set
Divided the data set into training and testing group
Classification:
for $i = 1$ to Z **do**
  Mapping functions (inputs)
  Update Augmented Weight Matrix(θ)
  Cost function or Average Cost J(θ)
  if $J(\theta) \leq \in \| \ (N == N_{max})$ **then**
    Optimum weights(θ)
  else
    Update Augmented Weight Matrix(θ)
  end if
end for
Find Normal CR-IoT users and Malicious Users
Plot the normal CR-IoT user data
Plot the malicious CR-IoT user data
Calculate the confusion matrix
Plot the confusion matrix value
Plot the AUC for LR

In SVM, the following optimization problem is defined which is maximising the classifier margin as follows:

$$\begin{aligned} \min \quad & \left(\frac{1}{2}\|w\|^2\right) \\ \text{s.t.} \quad & y_i\left(w.x_i + b\right) \geq 1 \end{aligned} \tag{11}$$

where $\|.\|$ represents the norm which is defined as $\|w\|^2 = w.w$, and b denotes the bias that shifting the hyperplane away from it's origin.

Now, the decision function of the hyperplane can then be written as follows:

$$g(x) = \delta\left(\sum_{i=1:\alpha_i>0} y_i\alpha_i\left(x.x_i\right) + b\right) \tag{12}$$

Algorithm 2 SVM based MUs classification ML algorithm in an energy harvested CR-IoT network.

Input: $\mathbb{X}$ N, M and L
Output: Normal CR-IoT users set and Malicious Users set
Initialisation :
Sensing the data
for $i = 1$ to N **do**
  for $t = l$ to L **do**
    Energy reported by the i^{th} CR-IoT users based on Eq. 3
  end for
end for
for $i = 1$ to M **do**
  for $t = l$ to L **do**
    Energy reported by the i^{th} malicious users based on Eq. 3
  end for
end for
Data processing for SVM:
Combining the both users data in to one data set
Find the number of features from data set
Extract the features matrix X and label vector Y
Scaling the features data set
Divided the data set into training and testing group
Classification:
Selected the linear support vector classifier (SVC)
Fit the data in the linear SVC
Find Normal CR-IoT users and Malicious Users
Plot the normal CR-IoT user data
Plot the malicious CR-IoT user data
Calculate the confusion matrix
Plot the confusion matrix value
Plot the AUC for SVM

where, $g(x)$ is the decision function of the hyperplane, $\sum$ is the sum of support vector/data, δ is the sign function, $x.x_i$ is the kernel function and the sum is over support vectors. The classification process compares the new instance x with each of the support vectors. Moreover, the $(x_i.x)$ measures how similar the new instance x is to the training instance x_i. Here, α_i are called support vectors that measure the contribution of x_i, i.e., α_i measures how important the given support vector and the decision function are defined by these vectors. We multiply by y_i to take into account the influence of the given support vector on the classification.

The SVM based MUs classification ML algorithm in an energy harvested CR-IoT network is shown in Algorithm 2 which consists of four stages: data generation, sensing, classification, and AUC.

C. k-Nearest Neighbors (k-NN) Algorithm

The k-NN algorithm is one of the easy-to-implement supervised ML algorithms. Although it is used in the solution of both regression and classification tasks, it used mainly in the solution of classification tasks in industry.

Algorithm 3 k-NN based MUs classification ML algorithm in an energy harvested CR-IoT network.

Input: D N, M and L
Output: Normal CR-IoT users set and Malicious Users set
Initialisation :
Sensing the data
for $i = 1$ to N **do**
  for $t = l$ to L **do**
    Energy reported by the i^{th} CR-IoT users based on Eq. 3
  end for
end for
for $i = 1$ to M **do**
  for $t = l$ to L **do**
    Energy reported by the i^{th} malicious users based on Eq. 3
  end for
end for
Data processing for k-NN:
Combining the both users data in to one data set
Find the number of features from data set
Extract the features matrix X and label vector Y
Scaling the features data set
Divided the data set into training and testing group
Classification:
Choose the value of K
for $i = 1$ to Z **do**
  Compute standard Euclidean distance, $d(x, y)$ between ith data point and each row of training data point by using the Eq. 13
  Sort the computed distances
  Find closest k-nearest neighbors point
  Vote for labels
end for
Find normal CR-IoT users and abnormal CR-IoT users (MUs)
Plot the normal CR-IoT user data
Plot the malicious CR-IoT user data
Calculate the confusion matrix
Plot the confusion matrix value
Plot the AUC for k-NN

In k-NN, the Euclidean distance is calculated based on the following equations:

$$d(x,y) = \sqrt{\sum_{i=1}^{n} c_i \left(x_i - y_i\right)^2} \tag{13}$$

where, c_i is the weight, and d represents distance between x_i and y_i.

The k-NN based MUs classification ML algorithm in an energy harvested CR-IoT network is shown in Algorithm 3 which consists of four stages: data generation, sensing, classification, and AUC.

D. Accuracy of the ML Algorithms

In this section, we can calculate the accuracy of ML algorithms as follows:

$$\eta = \frac{(TN + TP)}{(TP + FP + TN + FN)} \tag{14}$$

where, η is accuracy of the ML algorithms, and TP, FP, TN and FN indicates *True Positives*, *False Positives*, *True Positives* and *False Negatives*, respectively.

E. Global Decision

When the classification is done through the ML algorithms, all normal CR-IoT users execute local decision independently based on the sequential manner, and then transmit their decisions to the FC during the reporting time slot. After that, the FC makes a global decision about the PU activities like the absence and presence of the PU signal using the DS evidence theory under the sequential manner [2] as follows:

$$\begin{aligned} m\left(\mathbb{H}_0\right) &= \omega_1 m_1\left(\mathbb{H}_0\right) \oplus \omega_2 m_2\left(\mathbb{H}_0\right) \oplus \cdots \oplus \omega_N m_N\left(\mathbb{H}_0\right) \\ &= \frac{\sum_{\Gamma_1 \cap \Gamma_2 \cap \cdots \cap \Gamma_N = \mathbb{H}_0} \Pi_{i=1}^{N} \omega_i m_i\left(\Gamma_i\right)}{1 - \sum_{\Gamma_1 \cap \Gamma_2 \cap \cdots \cap \Gamma_N = \varnothing} \Pi_{i=1}^{N} \omega_i m_i\left(\Gamma_i\right)} \end{aligned} \tag{15}$$

$$\begin{aligned} m\left(\mathbb{H}_1\right) &= \omega_1 m_1\left(\mathbb{H}_1\right) \oplus \omega_2 m_2\left(\mathbb{H}_1\right) \oplus \cdots \oplus \omega_N m_N\left(\mathbb{H}_1\right) \\ &= \frac{\sum_{\Gamma_1 \cap \Gamma_2 \cap \cdots \cap \Gamma_N = \mathbb{H}_1} \Pi_{i=1}^{N} \omega_i m_i\left(\Gamma_i\right)}{1 - \sum_{\Gamma_1 \cap \Gamma_2 \cap \cdots \cap \Gamma_N = \varnothing} \Pi_{i=1}^{N} \omega_i m_i\left(\Gamma_i\right)} \end{aligned} \tag{16}$$

where

$$m_i\left(\mathbb{H}_0\right) = \int_{E_i}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_i\left(\mathbb{H}_0\right)} exp\left(-\frac{\left(\mathbb{X}_i - \mu_i\left(\mathbb{H}_0\right)\right)^2}{2\sigma_i^2\left(\mathbb{H}_0\right)}\right)$$

$$m_i\left(\mathbb{H}_1\right) = \int_{\infty}^{E_i} \frac{1}{\sqrt{2\pi}\sigma_i\left(\mathbb{H}_1\right)} exp\left(-\frac{\left(\mathbb{X}_i - \mu_i\left(\mathbb{H}_1\right)\right)^2}{2\sigma_i^2\left(\mathbb{H}_1\right)}\right)$$

here, $m_i\left(\mathbb{H}_0\right)$, $m_i\left(\mathbb{H}_1\right)$ and $m_i\left(\Omega\right)$ are the basic probability assignment (BPA) hypotheses of i^{th} normal CR-IoT users under $\mathbb{H}_0$, $\mathbb{H}_1$, and Ω, respectively. Moreover, Γ_i is an element of the set $\{\mathbb{H}_0, \mathbb{H}_1, \Omega\}$ and ω_i is the weight factor of the i^{th} CR-IoT user.

At the FC, the final combination result $m\left(\mathbb{H}_0\right)$ and $m\left(\mathbb{H}_1\right)$ by each normal CR-IoT user is obtained, then it makes a global decision $gd_{f|d}$ as follows:

$$gd_{f|d} = \begin{cases} 0; & \text{if } m\left(\mathbb{H}_0\right) > m\left(\mathbb{H}_1\right) \\ 1; & \text{if } m\left(\mathbb{H}_1\right) \geq m\left(\mathbb{H}_0\right) \end{cases} \tag{17}$$

F. Sum Rate Analysis

After calculating the detection performance at the FC based on the sequential manner in the previous subsection, now the sum rate is evaluated when considering numerous premises. During the transmission phase, the CR-IoT $\mathbb{T}x$ transmits its own relevant information towards the respective CR-IoT $\mathbb{R}x$ based on round robin scheduling approach [30]. In the case of the non-false alarm, if the

PU is absent, each unlicensed CR-IoT user is correctly sensed the absence of the PU; then each unlicensed CR-IoT user is likely to allow the licensed spectrum of the PU for a certain amount of time , as described by the probability of $(1 - gd_f)$. In another side, in the case of detection, the CR-IoT users should not interfere with the PU transmission. Consequently, the sum rate based on the round robin scheduling approach of the proposed ML algorithms is defined as follows:

$$R_{sum} = \rho g d_d C_{PU} + (1 - \rho)\, g d_f C_{i,CR-IoT} \tag{18}$$

where C_{PU} is the channel capacity of the PU link, $C_{i,CR-IoT}$ is the channel capacity of the i^{th} CR-IoT link, and $\rho \in [0, 1]$ indicates the primary activity factor which means the probability of the PUs transmitting in a given frame.

The C_{PU} and $C_{i,CR-IoT}$ are given as follows:

$$C_{PU} = log_2\,(1 + SNR_{PU}) \tag{19}$$

$$C_{i,CR-IoT} = \frac{T - \tau_s - \tau_r}{T} log_2\,(1 + SNR_{i,CR-IoT}) \tag{20}$$

SNR_{PU} is the SNR of the PUs link, $SNR_{i,CR-IoT}$ is the SNR of the i^{th} CR-IoT link, and T is the total frame length.

G. Network Lifetime Analysis

Now, we can measured the average energy consumption of the proposed ML algorithms in this section as follows:

$$E_{avg} = e_s \tau_s + e_t T_T \left(\rho g d_d + (1 - g d_f)\,(1 - \rho)\right) \tag{21}$$

where e_s denotes the energy consumed for the sensing duration, T_T denotes the transmission time, i.e., $T_T = T - \tau_s - \tau_r$, and τ_s, e_t denotes the energy consumed for the transmission duration,

The expected network lifetime (ξ) of the proposed ML algorithm can be calculated as follows:

$$\xi = \frac{e_c + e_s^h}{E_{avg}} \tag{22}$$

where e_c is the capacity of battery and e_s^h is the energy harvested during the sensing phase (the flexible sensing time duration).

V. Simulation Results and Discussions

In this section, we validate the theoretical findings, and analyze the detection performance of the proposed the ML algorithms. This is achieved through numerical simulations via Matlab. Monte-Carlo simulations were carried out using the simulation parameters listed in Table II below which are based on the rationale of the other researchers [26], [27], [31]. The performance of the proposed scheme based on the ML algorithm is compared with similar schemes, such as SVM based classification of MUs in CRNs [27], the detection of malicious PU emulation based on a SVM for a mobile CRN using software-defined radio [26], and CSS algorithm based on SVM against spectrum sensing-data-falsification (SSDF) attack [31].

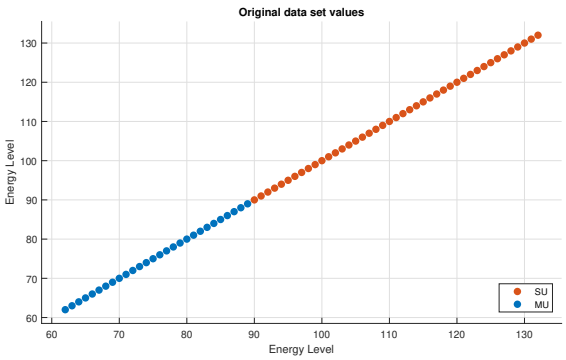

Fig. 3. Data set values.

TABLE II
Simulation parameters with values

Parameters	Value
The number of the CR-IoT users	43
The number of malicious users	28
The sampling frequency, f_s	300kHz
The sensing time slot, τ_s	1 ms
The reporting time slot, τ_r	1 ms
The time slot length, T	10 ms
The energy consumption in sensing phase, e_s	1 J
The energy consumption in transmission phase, e_t	3 J
The capacity of battery, e_c	301 J
The probability of the absence of the PU, ρ	0.5
The probability of the presence of the PU, $(1 - \rho)$	0.5

The data set consists of MUs and CR-IoT users/SUs classes categorized according to energy level. This dataset, consisting of 71 data, was first optimized and then training tests were passed. Fig. 3 shows the MU and CR-IoT user distribution chart in our data set.

After the data set was trained with classifiers, the accuracy scores of the obtained models were compared. The performances were checked using class predictions, confusion matrix, and Receiver Operating Characteristic (ROC) curve. The validation accuracy score predicts the performance of a model on new data compared to training data. The best model was chosen based on this score. k-fold cross-validation was used to calculate the accuracy points using the observations in k validation folds and to determine the mean cross-correct error and was considered as 5. It calculates the confusion matrix and ROC curve based on these estimates by making predictions on k-fold cross-validation observations.

In Fig. 4, the LR prediction values are shown. SU and MU correct predictions are given with circles and the miss predictions are given with x shape. Also in Fig. 5, the confusion matrix are shown. The confusion matrix is the most important metric commonly used to evaluate classification patterns. Some metrics can be derived from the confusion

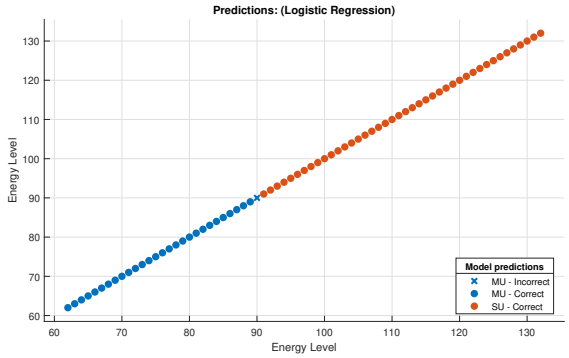

Fig. 4. The prediction value of the LR algorithm.

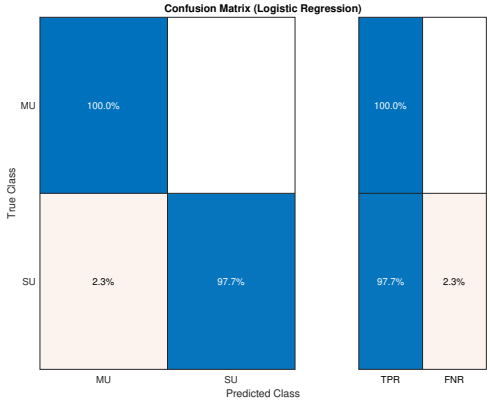

Fig. 5. The confusion matrix of the LR algorithm.

matrix like accuracy. Accuracy is the proposed model's overall predicted accuracy and it can be formalized using the Eq. 14.

As can be seen from confusion matrix, *True Positive Rate (TPR)* indicates positive values that have been correctly predicted and *False Positive Rate (FPR)* shows negative values that have been incorrectly predicted. *True Negative Rate (TNR)* gives negative values that have been correctly predicted and *False Negative Rate (FNR)* indicates positive values that have been incorrectly predicted.

The results of LR, SVM, KNN classifiers are given in the scatter plot given in Fig. 4, Fig. 7 and Fig. 10. After the classifiers are trained, the estimates of the models are shown with the help of these distribution charts. When these three figures are compared, the best model is LR with its predictions.

The confusion matrix plot was used to understand how the classifiers performed in each class and in which areas they performed poorly. In these graphs, the rows show the actual class and the columns the estimated class. In Fig. 5, Fig. 8 and Fig. 11 , confusion matrices obtained for LR, SVM and k-NN are given.

In Fig. 6, the ROC of the LR is given in detail. The ROC determines the correctness of a classification model at a user-defined threshold. As can be seen from Fig. 6, AUC gives the model's accuracy by ROC. The aim of the algorithm is to push the line towards one and maximize

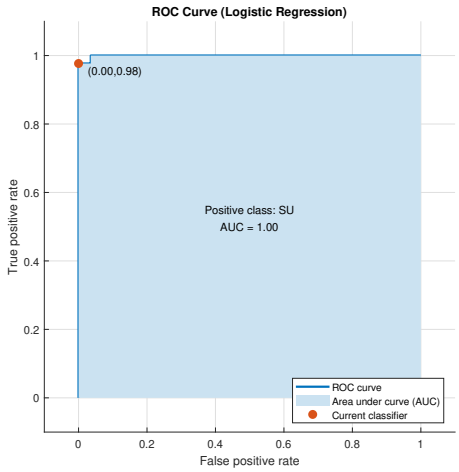

Fig. 6. ROC of the LR algorithm.

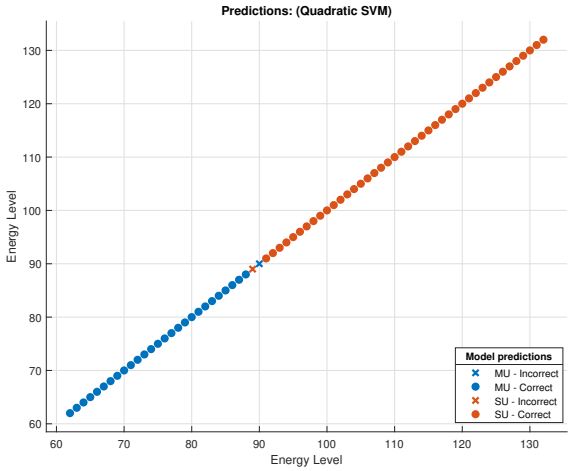

Fig. 7. The prediction value of the SVM algorithm.

the under the curve.

To see how classifiers perform per class, *TPR* and *FNR* are given. *TPR* expresses the proportion of correctly classified observations per actual class, while *FNR* indicates the proportion of misclassified observations per actual class.

The SVM prediction values are shown in Fig. 7. The means of the points are discussed for LR, and it can be seen from Fig. 7, the wrong predictions are increased in SVM. The confusion matrix is also given in Fig. 8, values of some areas are decreased in SVM confusion matrix. When these results are compared, LR obtained 100% to 97.7%, SVM 96.4% to 97.7%, and k-NN 100% to 81.4% in terms of determining MU and CR-IoT user, respectively.

In the light of these results, it has been observed that the LR model is more successful. In Fig. 6, Fig. 9 and Fig. 12, the ROC curves are given for LR, SVM and k-NN which show the correct and false positive rates after the models are trained. Here, true positive and false positive rates for the trained classifier are shown.

As Fig. 11 the k-NN confusion matrix shows the afore-

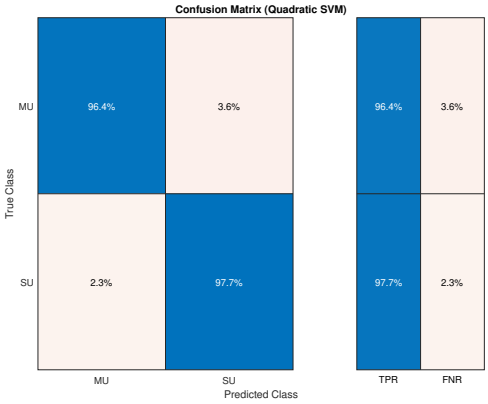

Fig. 8. The confusion matrix of of the SVM algorithm.

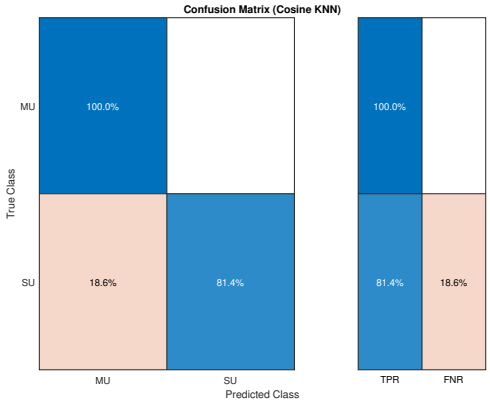

Fig. 11. The confusion matirx of the k-NN algorithm.

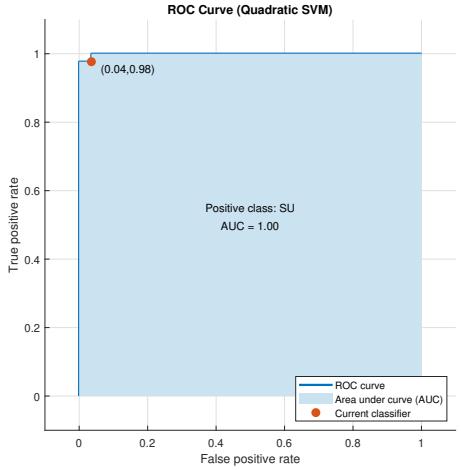

Fig. 9. ROC curve of the SVM algorithm.

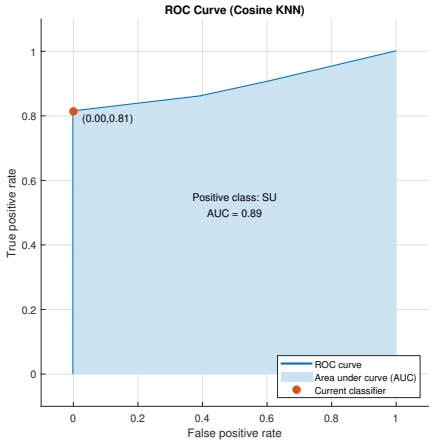

Fig. 12. ROC curve of the k-NN algorithm.

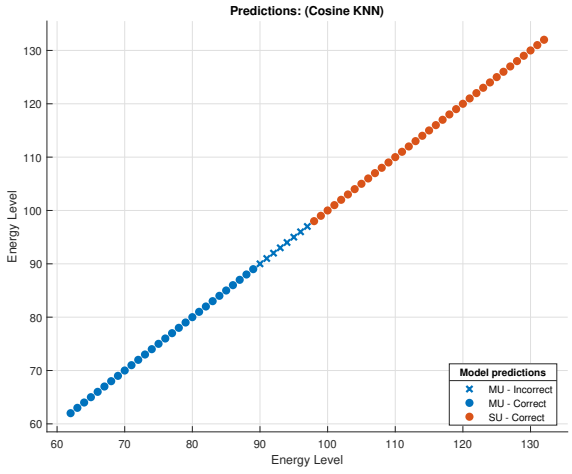

Fig. 10. The prediction value of the k-NN algorithm.

mentioned areas for accuracy calculation.

ROC curve of the k-NN algorithm is given in Fig. 12, and the AUC is reduced as 0.89 with positive class of SU. The k-NN is the most unsuccessful classification algorithm for the proposed system.

Fig. 13 demonstrates the sum rate of the ML algorithms (LR, SVM, k-NN). The sum rate is a function of the probability of false alarm. For example, the sum rate of the proposed LR algorithm is $3.45Hz$ compared to other SVM algorithm and the k-NN algorithm are $3.38Hz$ and $3.25Hz$, respectively when the probability of false alarm is 0.3. Therefore, the proposed LR algorithm is an enhanced sum rate when compared to both the SVM algorithm and the k-NN algorithm.

The energy consumption for the ML algorithms (LR, SVM, k-NN) shows in Fig. 14. The average energy consumption is a function of the probability of false alarm. The energy consumption of the proposed ML algorithms with the LR algorithm achieved a better energy efficient when compared to other ML algorithms because of its higher detection performance. For example, the energy consumption of the LR algorithm is $1.29J$ compared to other SVM algorithm and the k-NN algorithm are $1.33J$ and $1.49J$, respectively when the probability of false alarm is 0.2. Therefore, the proposed LR algorithm is a better than other ML algorithms with SVM and k-NN for any value of the probability of false alarm.

The expected lifetime of the proposed ML algorithms with the LR, SVM, and k-NN shows in Fig. 15. It is

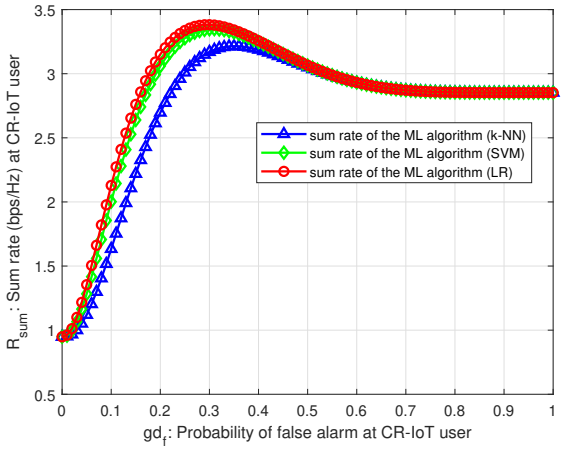

Fig. 13. The sum rate of the ML algorithms (k-NN, SVM, LR).

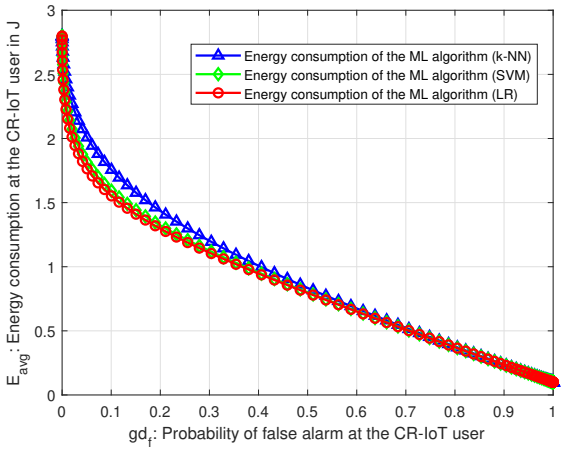

Fig. 14. The energy consumption of the ML algorithms (k-NN, SVM, LR).

clearly seen that the expected lifetime of the proposed ML algorithms with the LR algorithm has extended when compared to both SVM and k-NN algorithms. For example, the sum rate of the LR algorithm is $3.48Hz$ compared to other SVM algorithm and the k-NN algorithm are $3.37Hz$ and $3.28Hz$, respectively; when the expected life time at the CR-IoT user in round is 400. Moreover, with regard to the expected lifetime, the sum rate of the proposed ML algorithms are decreased when the expected lifetime in round is greater than 400. Moreover, the sum rate of the proposed ML algorithm is a same when the expected lifetime in round is greater than 850. Therefore, the proposed LR algorithm is a prolonged the expected lifetime when compared to both the SVM algorithm and the k-NN algorithm.

Accuracy scores for LR, SVM, and k-NN are given in Table III. In the light of these values, the LR model was preferred for the architecture we recommend.

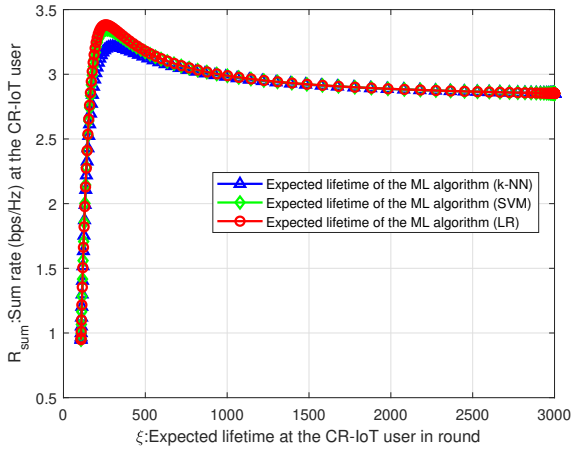

Fig. 15. The expected lifetime of the ML algorithms (k-NN, SVM, LR).

TABLE III
ACCURACY OF THE MACHINE LEARNING ALGORITHMS

Machine learning algorithms	Accuracy
LR	98.60%
SVM	97.20%
k-NN	60.60%

VI. Conclusion

Malicious user detection is very important issue for efficient use of spectrum in CR-IoT networks. With using improved malicious user detection schemes, more robust spectrum utilization will be provided. For detecting malicious users in CR-IoT networks, classification algorithms can be used. For this reason, we utilized machine learning based classification algorithms such as LR, SVM, and k-NN. And also, the proposed schemes have been applied in CR-IoT networks, and enhanced results have been obtained in terms of sensing gain, sum rate, and network lifetime compared to the conventional schemes. For future works, various datasets and classification algorithms can be considered for different CR-IoT networks scenarios.

Acknowledgment

This research was supported in part by the Islamic University (IU), Kushtia-7003, Bangladesh (Ref. No. 141/EDU/IU-2020/634 and by the Department of Information and Communication Technology, Islamic University, Kushtia, Bangladesh.

References

[1] S. Chen, H. Xu, D. Liu, B. Hu, and H. Wang, "A vision of iot: Applications, challenges, and opportunities with china perspective," *IEEE Internet of Things journal*, vol. 1, no. 4, pp. 349–359, 2014.

[2] M. S. Miah, "An energy efficient spectrum sensing scheme for the cognitive radio based internet of things," Ph.D. dissertation, NUI Galway, 2020.

[3] H. Afzal, M. Rafiq Mufti, A. Raza, and A. Hassan, "Performance analysis of qos in iot based cognitive radio ad hoc network," *Concurrency and Computation: Practice and Experience*, p. e5853.
[4] M. A. Hossain, M. Schukat, and E. Barrett, "Enhancing the spectrum utilization in cellular mobile networks by using cognitive radio technology," in *2019 30th Irish Signals and Systems Conference (ISSC)*. IEEE, 2019, pp. 1–6.
[5] M. R. Amin, M. M. Rahman, M. A. Hossain, M. K. Islam, K. M. Ahmed, B. C. Singh, and M. S. Miah, "Unscented kalman filter based on spectrum sensing in a cognitive radio network using an adaptive fuzzy system," *Big Data and Cognitive Computing*, vol. 2, no. 4, p. 39, 2018.
[6] J. Mitola, "Cognitive radio for flexible mobile multimedia communications," in *1999 IEEE International Workshop on Mobile Multimedia Communications (MoMuC'99)(Cat. No. 99EX384)*. IEEE, 1999, pp. 3–10.
[7] A. A. Khan, M. H. Rehmani, and A. Rachedi, "Cognitive-radio-based internet of things: Applications, architectures, spectrum related functionalities, and future research directions," *IEEE wireless communications*, vol. 24, no. 3, pp. 17–25, 2017.
[8] M. A. Hossain, M. Schukat, and E. Barrett, "Enhancing the spectrum sensing performance of cluster-based cooperative cognitive radio networks via sequential multiple reporting channels," *Wireless Personal Communications*, pp. 1–23, 2020.
[9] A. A. Khan, M. H. Rehmani, and A. Rachedi, "When cognitive radio meets the internet of things?" in *2016 international wireless communications and mobile computing conference (IWCMC)*. IEEE, 2016, pp. 469–474.
[10] A. Ali, L. Feng, A. K. Bashir, S. H. A. El-Sappagh, S. H. Ahmed, M. Iqbal, and G. Raja, "Quality of service provisioning for heterogeneous services in cognitive radio-enabled internet of things," *IEEE Transactions on Network Science and Engineering*, 2018.
[11] D. T. Otermat, I. Kostanic, and C. E. Otero, "Analysis of the fm radio spectrum for secondary licensing of low-power short-range cognitive internet of things devices," *IEEE Access*, vol. 4, pp. 6681–6691, 2016.
[12] Y. B. Zikria, F. Ishmanov, M. K. Afzal, S. W. Kim, S. Y. Nam, and H. Yu, "Opportunistic channel selection mac protocol for cognitive radio ad hoc sensor networks in the internet of things," *Sustainable Computing: Informatics and Systems*, vol. 18, pp. 112–120, 2018.
[13] Y. Gu, H. Chen, C. Zhai, Y. Li, and B. Vucetic, "Minimizing age of information in cognitive radio-based iot systems: Underlay or overlay?" *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 273–10 288, 2019.
[14] H. A. B. Salameh, S. Al-Masri, E. Benkhelifa, and J. Lloret, "Spectrum assignment in hardware-constrained cognitive radio iot networks under varying channel-quality conditions," *IEEE Access*, vol. 7, pp. 42 816–42 825, 2019.
[15] A. Sultan, "Sensing and transmit energy optimization for an energy harvesting cognitive radio," *IEEE wireless communications letters*, vol. 1, no. 5, pp. 500–503, 2012.
[16] H. A. Shah and I. Koo, "Reliable machine learning based spectrum sensing in cognitive radio networks," *Wireless Communications and Mobile Computing*, vol. 2018, 2018.
[17] S. U. Jan, V.-H. Vu, and I. Koo, "Throughput maximization using an svm for multi-class hypothesis-based spectrum sensing in cognitive radio," *Applied Sciences*, vol. 8, no. 3, p. 421, 2018.
[18] J. Zhu, Y. Song, D. Jiang, and H. Song, "A new deep-q-learning-based transmission scheduling mechanism for the cognitive internet of things," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 2375–2385, 2017.
[19] M. A. Rahman, Y.-D. Lee, and I. Koo, "An efficient transmission mode selection based on reinforcement learning for cooperative cognitive radio networks," *Human-centric Computing and Information Sciences*, vol. 6, no. 1, p. 2, 2016.
[20] K. M. Thilina, K. W. Choi, N. Saquib, and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks," *IEEE Journal on selected areas in communications*, vol. 31, no. 11, pp. 2209–2221, 2013.
[21] M. Alshawaqfeh, X. Wang, A. R. Ekti, M. Z. Shakir, K. Qaraqe, and E. Serpedin, "A survey of machine learning algorithms and their applications in cognitive radio," in *International Conference on Cognitive Radio Oriented Wireless Networks*. Springer, 2015, pp. 790–801.
[22] Y.-D. Huang, Y.-C. Liang, and G. Yang, "A fuzzy support vector machine algorithm for cooperative spectrum sensing with noise uncertainty," in *2016 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2016, pp. 1–6.
[23] Z. Li, W. Wu, X. Liu, and P. Qi, "Improved cooperative spectrum sensing model based on machine learning for cognitive radio networks," *IET Communications*, vol. 12, no. 19, pp. 2485–2492, 2018.
[24] A. Albehadili, A. Ali, F. Jahan, A. Y. Javaid, J. Oluochy, and V. Devabhaktuniz, "Machine learning-based primary user emulation attack detection in cognitive radio networks using pattern described link-signature (pdls)," in *2019 Wireless Telecommunications Symposium (WTS)*. IEEE, 2019, pp. 1–7.
[25] S. Srinivasan, K. Shivakumar, and M. Mohammad, "Semi-supervised machine learning for primary user emulation attack detection and prevention through core-based analytics for cognitive radio networks," *International Journal of Distributed Sensor Networks*, vol. 15, no. 9, p. 1550147719860365, 2019.
[26] E. Cadena Muñoz, L. F. Pedraza Martínez, and J. E. Ortiz Triviño, "Detection of malicious primary user emulation based on a support vector machine for a mobile cognitive radio network using software-defined radio," *Electronics*, vol. 9, no. 8, p. 1282, 2020.
[27] M. S. Khan, L. Khan, N. Gul, M. Amir, J. Kim, and S. M. Kim, "Support vector machine-based classification of malicious users in cognitive radio networks," *Wireless Communications and Mobile Computing*, vol. 2020, 2020.
[28] M. S. Khan, N. Gul, J. Kim, I. M. Qureshi, and S. M. Kim, "A genetic algorithm-based soft decision fusion scheme in cognitive iot networks with malicious users," *Wireless Communications and Mobile Computing*, vol. 2020, 2020.
[29] Y. Zhang, Q. Wu, and M. Shikh-Bahaei, "Ensemble learning based robust cooperative sensing in full-duplex cognitive radio networks," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2020, pp. 1–6.
[30] M. S. Miah, M. Schukat, and E. Barrett, "Sensing and throughput analysis of a mu-mimo based cognitive radio scheme for the internet of things," *Computer Communications*, 2020.
[31] H. Zhu, T. Song, J. Wu, X. Li, and J. Hu, "Cooperative spectrum sensing algorithm based on support vector machine against ssdf attack," in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2018, pp. 1–6.

Dr. Md. Sipon Miah received his B.Sc., M.Sc. and Ph.D. in Information and Communication Technology (ICT) from the Islamic University (IU), Kushtia-7003, Bangladesh, in 2006, 2007 and 2016, respectively. Dr. Sipon received a Structured Ph.D. in the School of Computer Science from the National University of Ireland Galway (NUIG), Galway, Ireland, in 2020. In 2013, Dr. Sipon was awarded the prestigious ICT Scholarship (Bangladesh). In 2016, Dr. Sipon was awarded the prestigious Hardiman Scholarship (Ireland). Since 2010, he has been with the Department of Information and Communication Technology (ICT), in the Islamic University (IU), Kushtia-7003, Bangladesh. He is currently an Associate Professor in the same Department. His research interests include Spectrum Sensing, Energy Harvesting, Internet of Things, MU-MIMO based Cognitive Radio Networks, Massive MIMO based Cognitive Radio Networks and Machine Learning based Cognitive Radio based IoT.

Mohammad Amzad Hossain received his B.Sc., and M.Sc. in Information and Communication Technology (ICT) from the Islamic University (IU), Kushtia, Bangladesh, in 2010 and 2011, respectively. He is currently an Assistant Professor in the Department of Information and Communication Engineering (ICE), Noakhali Science and Technology University, Noakhali, Bangladesh. Amzad is currently pursuing a Structured Ph.D. in the School of Computer Science, National University of Ireland Galway (NUIG), Galway, Ireland. In 2018, Amzad has awarded the prestigious College of Science and Engineering postgraduate research Scholarship. His research interests include Spectrum Sensing, MIMO based Cognitive Radio Networks, Cognitive Radio based Internet of Things (CR-IoT) Networks and Deep Learning.

Kazi Mowdud Ahmed received his B.Sc. (Hon's) and M.Sc. in Information and Communication Technology (ICT) from the Islamic University (IU), Kushtia-7003, Bangladesh, in 2012 and 2013, respectively. Since 2018, he has been with the Department of Information and Communication Technology (ICT), in the Islamic University (IU), Kushtia-7003, Bangladesh. He is currently a lecturer in the same department. He worked as a research assistant in Multimedia Communication Systems Lab. (MCSL) and Computer Vision and Intelligent Interfacing Lab.(CVIIL) from 2011 to 2016 in the same department. Since 2018, Mowdud is working as a research assistant in Wireless Communications Lab (WCL) in the department of Information and Communication Technology, Islamic University, Kushtia-7003, Bangladesh. His research interests includes Cloud Computing, Machine Learning and Deep Learning, AI-enabled Cognitive Radio based IoT.

Dr. Md. Mahbubur Rahman is a Professor, Department of Information and Communication Technology, Islamic University, Kushtia, Bangladesh. He received the B.Sc. and M.Sc. degrees in Physics, Rajshahi University, Rajshahi, Bangladesh. In 1997, Rahman received his Ph.D. in Computer Science & Engineering from Rajshahi University, Rajshahi, Bangladesh. He worked as a dean, faculty of Applied Science and Technology, Islamic University, Kushtia, Bangladesh. Since 1998, Mabubur is working as a director in Wireless Communications Laboratory (WCL) in the department of Information and Communication Technology, Islamic University, Kushtia-7003, Bangladesh. He has published fifty reputed journal papers. His main research interests include mobile communications, wireless sensor networks, Internet of things, cognitive radio networks and AI-enabled networking.

Dr. Ali Calhan received his M.Sc. and Ph.D. degrees from the University of Kocaeli, Turkey in 2006 and 2011. Science 2011, he has been a member of the Computer Engineering Department of Duzce University. Currently, he is an Associate Professor in the Department of Computer Engineering, Duzce University (DU). His research interests are wireless communications, cognitive radio networks, body area networks and software-defined networks.

Dr. Murtaza Cicioglu received his Ph.D. degree in Electrical-Electronic and Computer Engineering from Düzce University, Turkey in 2020. Since 2009, he has been a member of the Information Technologies Department of Ministry of National Education, Turkey. He is a member of the Software Defined Networks Community, IEEE. His research interests include software-defined networking, wireless communications, 5G, body area networks, cognitive radio, and Riverbed Modeler (OPNET) simulation software.