

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2021.Doi Number

# Experimental Setup for Online Fault Diagnosis of Induction Machines via Promising IoT and Machine Learning: Towards Industry 4.0 Empowerment

Minh-Quang Tran<sup>1,2</sup>, Mahmoud Elsisy<sup>1,3</sup>, Karar Mahmoud<sup>4,5</sup>, Meng-Kun Liu<sup>1,6</sup>,  
Matti Lehtonen<sup>4</sup>, and Mohamed M. F. Darwish<sup>3,4,\*</sup>

<sup>1</sup>Industry 4.0 Implementation Center, Center for Cyber-physical System Innovation, National Taiwan University of Science and Technology, 10607 Taipei, Taiwan.

<sup>2</sup>Department of Mechanical Engineering, Thai Nguyen University of Technology, 3/2 Street, Tich Luong ward, 250000 Thai Nguyen, Vietnam.

<sup>3</sup>Department of Electrical Engineering, Faculty of Engineering at Shoubra, Benha University, 11629 Cairo, Egypt.

<sup>4</sup>Department of Electrical Engineering and Automation, School of Electrical Engineering, Aalto University, 02150 Espoo, Finland.

<sup>5</sup>Department of Electrical Engineering, Faculty of Engineering, Aswan University, 81542 Aswan, Egypt.

<sup>6</sup>Department of Mechanical Engineering, National Taiwan University of Science and Technology, Taipei, 10607, Taiwan.

Corresponding author: M. M. F. Darwish (e-mail: mohamed.m.darwish@aalto.fi; mohamed.darwish@feng.bu.edu.eg)

This work was supported in part by the Department of Electrical Engineering and Automation, Aalto University, Espoo, Finland, and in part by the Center for Cyber-physical System Innovation from the Featured Areas Research Center Program in the Agenda of the Higher Education Sprout Project, Taiwan.

**ABSTRACT** In recent years, the internet of things (IoT) represents the main core of Industry 4.0 for cyber-physic systems (CPS) in order to improve the industrial environment. Accordingly, the application of IoT and CPS has been expanded in applied electrical systems and machines. However, cybersecurity represents the main challenge of the implementation of IoT against cyber-attacks. In this regard, this paper proposes a new IoT architecture based on utilizing machine learning techniques to suppress cyber-attacks for providing reliable and secure online monitoring for the induction motor status. In particular, advanced machine learning techniques are utilized here to detect cyber-attacks and motor status with high accuracy. The proposed infrastructure validates the motor status via communication channels and the internet connection with economical cost and less effort on connecting various networks. For this purpose, the CONTACT Element platform for IoT is adopted to visualize the processed data based on machine learning techniques through a graphical dashboard. Once the cyber-attacks signal has been detected, the proposed IoT platform based on machine learning will be visualized automatically as fake data on the dashboard of the IoT platform. Different experimental scenarios with data acquisition are carried out to emphasize the performance of the suggested IoT topology. The results confirm that the proposed IoT architecture based on the machine learning technique can effectively visualize all faults of the motor status as well as the cyber-attacks on the networks. Moreover, all faults of the motor status and the fake data, due to the cyber-attacks, are successfully recognized and visualized on the dashboard of the proposed IoT platform with high accuracy and more clarified visualization, thereby contributing to enhancing the decision-making about the motor status. Furthermore, the introduced IoT architecture with Random Forest algorithm provides an effective detection for the faults on motor due to the vibration under industrial conditions with excellent accuracy of 99.03% that is significantly greater than the other machine learning algorithms. Besides, the proposed IoT has low latency to recognize the motor faults and cyber-attacks to present them in the main dashboard of the IoT platform.

**INDEX TERMS** Fault diagnosis; induction motor; machine learning; internet of things; Industry 4.0.

## I. INTRODUCTION

Worldwide, since the industrial revolution, rotary equipment has been widely used in many areas. Among them, induction motors play an essential role in the industry due to their low cost, easy maintenance, and unrestricted working environment [1]. However, if the production line is unpredictably shut down because of sudden failures of motors, it causes heavy damage over time. Therefore, fault detection and diagnosis for induction motors become much more important. The motor failures would generate heat and vibration and increase energy consumption [2]. As a result, the motor life would be reduced. There are many kinds of motor failures which are categorized into two major types of faults including electrical faults and mechanical faults. Mechanical faults include bearing faults, misalignment faults, and air-gap eccentricity faults. While electrical faults are related to faults like stator short-circuits, broken bar faults, and end ring faults [3, 4]. Bearings are among the most important components in induction machines. Their load capability, running accuracy, noise levels, frictional heat, and useful life will directly affect the induction machines [5]. It is about 40% - 90% of all motor failures that come from bearing faults depending on the size of the motors [6]. Therefore, the early detection and diagnosis of bearing faults could prevent sudden failure. This research subject still attracts great attention from the research community.

The diagnosis method can be divided into vibration and current analysis. The approach based on the motor current signature analysis (MCSA) is considered because of its low-cost, non-intrusive, and fast installation [7, 8]. However, the current signal diagnosis is difficult to detect abnormalities in the early stage of motor failures, this approach is still an immature technology compared to the vibration signal diagnosis [9]. According to the specification of ISO-10816 [10], it proposed a method for evaluating the statuses of the motor based on the measured vibration signal. A threshold for vibration damage has been proposed for motor condition monitoring. The method based on the curve component analysis was proposed to analyze vibration characteristics and to establish nonlinear data for the training dataset [11]. The individual frequency bands for envelope spectrum analysis can be extracted by filtering the vibration signal with the appropriate bandwidth. Then, the characteristic frequencies of the vibration signals can be observed in the initial stage of the bearing fault [12]. Statistical calculations were used to calculate the measured vibration energy. The influence of bearing speed on statistical indicators has been introduced, it showed that the kurtosis and skewness factors can be effective parameters for bearing defect identification [13].

Signal analysis approaches have been widely used to study motor signal vibration for bearing fault diagnosis. An improved fast Fourier transform for vibration signal analysis was introduced for both simulation and practical applications [14]. Several time-frequency analysis approaches were utilized for fault diagnoses such as the short-time Fourier Transform (STFT), Wavelet Transform (WT), and the Hilbert-

Huang Transform (HHT). The WT could provide a good resolution for both time and frequency domain that overcome the limitation of the STFT, thus the method is widely used for fault diagnosis [15, 16]. In addition to Fourier transform and WT, the HHT is also an effective approach for motor vibration signal analysis, it can detect malfunctioning by revealing the instantaneous amplitude and nonlinear and nonstationary characteristics in the frequency content [17]. Furthermore, the vibration signal contaminated by signal noise could be filtered by using the wavelet packet decomposition (WPD), the energy of the WPD coefficients could be utilized to detect the rolling bearing failures effectively [18]. The WPD and statistical methods were developed to successfully extract the significant features of bearing faults using time-frequency analysis [19].

Recently, fault diagnosis and detection based on artificial machine learning techniques have been performed effectively [20]. These methods usually involve intelligent signal analysis and it requires data collection to do signal processing and feature extraction. Various classifiers including artificial neural networks, k-nearest neighbors (kNN), and support vector machines (SVM) have been applied to fuse the vibration and current signal for multi-faults classification of induction motor [9]. Yang et al. [21] proposed a method based on a random forest (RF) algorithm that can achieve satisfactory accuracy for machine fault diagnosis with fast execution speed. The performance of the proposed method is much better compared to the other classifiers such as kNN, SVM, and decision tree. The machine learning ensemble classifier including RF and extreme gradient boosting (XGBoost) could be a promising model in terms of accuracy for bearing fault diagnosis [22, 23]. However, it is still difficult to determine features that are the most sensitive to the fault of the induction machine, because the performance of machine learning techniques is reflected by choosing suitable fault features. Deep learning can extract and learn the representative patterns of the signals effectively compared to conventional feature extraction and selection methods. For example, the convolutional neural network method shows excellent performance for pattern recognition that was also applied for fault diagnosis [24, 25]. Janssens et al. developed an architecture of the CNN model for detecting rotary machinery faults with vibration spectrum features [26]. However, the computational time is one of the most challenging to apply deep learning for online condition monitoring. Thus, an appropriate machine learning approach still needs to investigate to enhance the performance of the model with high accuracy and short computational time. The authors of [27] have investigated the cloud-based malware detection game, where mobile devices divest their application traces to security servers through base stations or access points.

Nowadays, the development of Industry 4.0 and the rise of the Internet of things (IoT) have become a solution for advanced predictive maintenance applications, in which the status of the motor is constantly monitored and recorded [28].

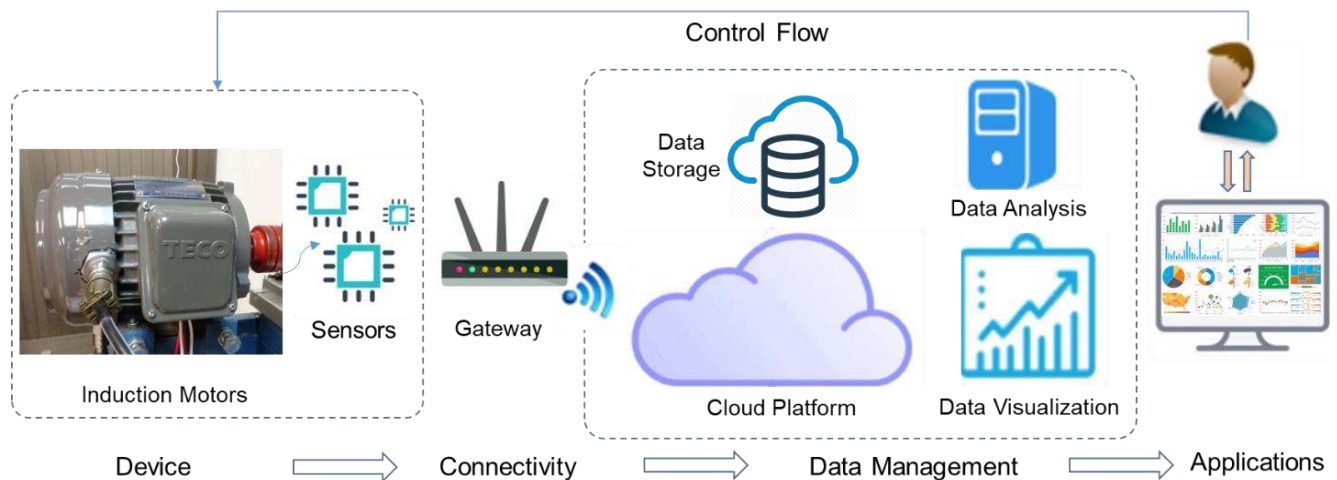


FIGURE 1. Industrial IoT architecture for condition monitoring of induction motors.

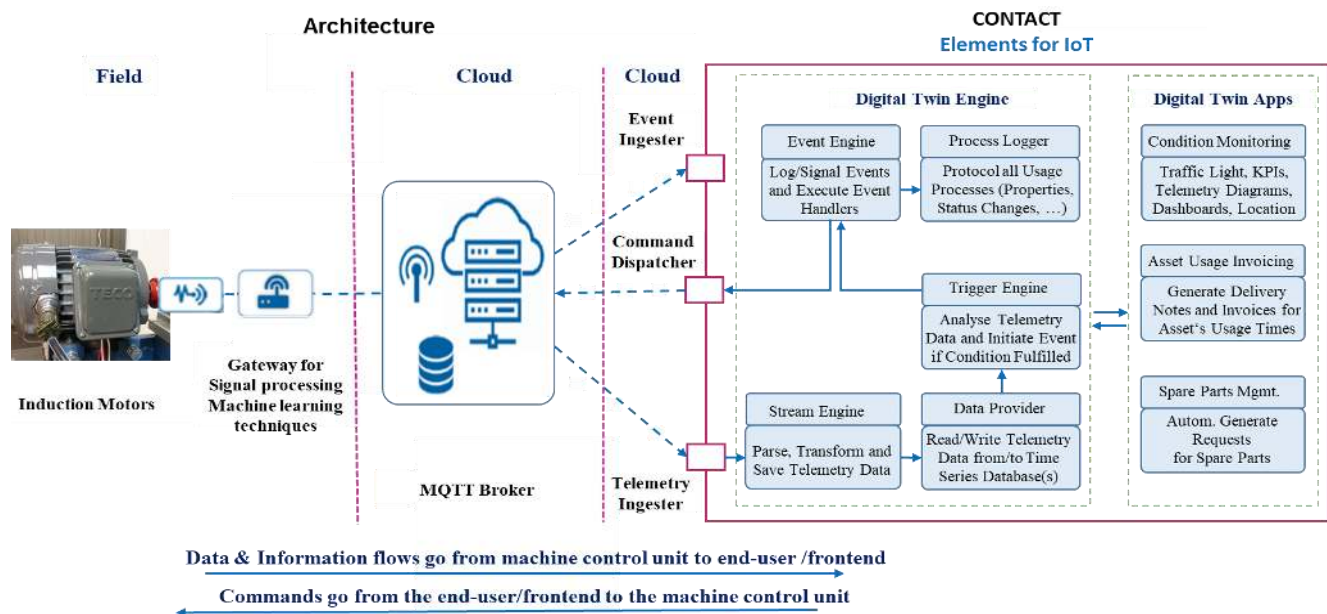


FIGURE 2. Proposed IoT architecture for online monitoring of the induction motor.

The concept of industrial IoT has been considered in many industrial fields and is expected to drive industry growth. The industrial IoT concept involves smart machines that are equipped with different sensors, connectivity, data management, and application layer as shown in Fig. 1 [29]. The sensor data analytics and management of big data are both challenging tasks of IoT, particularly for online condition monitoring of the IoT system [30]. IoT system allows the user able to monitor the machine status in real-time and it can give quick feedback with proper adjustments [31]. Thus it can help to accelerate the operation of the factories that involve diverse electrical machines, mostly induction motors.

As aforementioned in the literature, effective monitoring bearing faults are considered the main challenges of the rotary equipment to enhance the efficiency of the operation and lifetime of the induction machines. This study was inspired by the development of IoT and machine learning approaches that towards intelligence industrial machine. To the best of the authors' knowledge, this is the first study to experimentally apply IoT and CPS concepts to effectively diagnose induction motor faults with data acquisition. Specifically, an intelligence IoT platform integrating an improved machine learning algorithm is developed to detect the bearing fault of the induction motor. Besides, the powerful learning ability of the

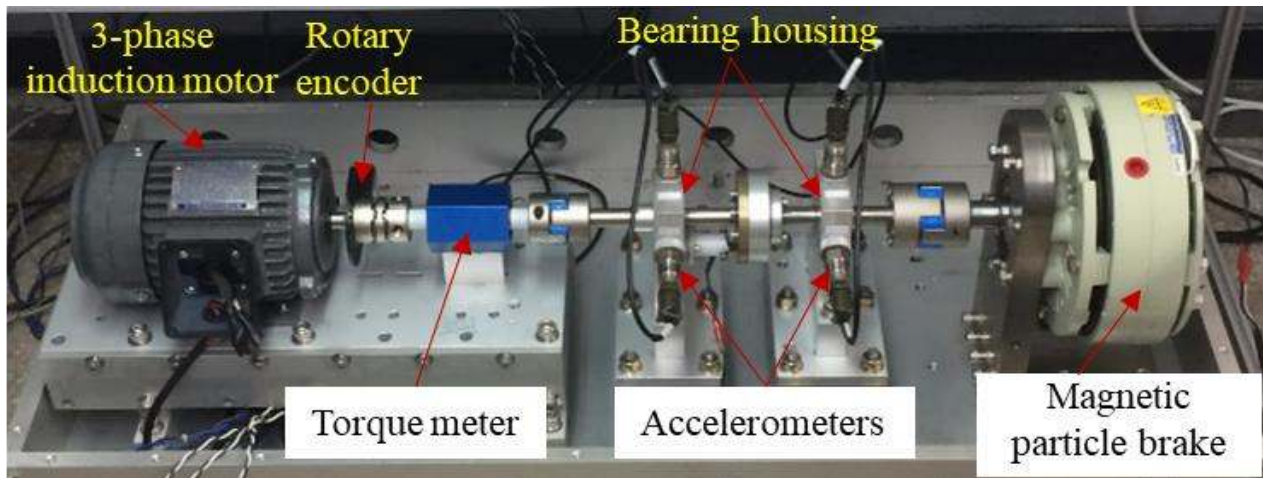


FIGURE 3. Experimental motor-driven rotary system setup.

machine learning model enables the proposed system to detect fake data considered as cyber-attack accurately.

The contributions of this paper are summarized in the following points,

- Developing a cheap and easy implementation IoT architecture with machine learning techniques for motor faults detection instead of the traditional faults detection schemes for induction machines.
- The introduced IoT architecture with machine learning algorithms can also detect the cyber-attack and present it as fake data in the main dashboard of the IoT platform to provide secure online monitoring.
- The proposed infrastructure can detect the normal, inner bearing defect, outer bearing defect in order to maintain a healthy state, and increase the lifetime of the induction machines.
- The proposed IoT has low latency to recognize the motor faults and cyber-attacks to present them in the main dashboard of the IoT platform.
- The experimental results indicate the superiority of the proposed IoT system integrating machine learning for condition monitoring of the induction machines.

The remainder of the paper is organized as follows. The proposed industrial IoT architecture and machine learning algorithms are described in Sections II and III. Simulation results are presented in Section IV. The conclusions drawn from the result section are summarized in Section V.

## II. PROPOSED INDUSTRIAL IOT ARCHITECTURE

### A. ARCHITECTURE OVERVIEW

The extension of the IoT in industrial applications is namely industrial internet of things (IIoT) which refers to the communication of machine-to-machine (M2M), data analysis, and machine learning. Recently, Industrial IoT has been considered in many applications towards smart factories to get high performance, more efficiency, and reliability in their

operations [32]. An architecture of IIoT for condition monitoring of induction motors is proposed in Fig. 2. The induction machines are equipped with sensors to continuously collect the data. Then, the sensor data is transmitted on the cloud layer through the connectivity layer with the IoT gateway. To manage the sensor data, the software part is required, then the large volumes of sensor data can be stored and analyzed in real-time. Finally, the decision-making is carried out before any necessary actions feedback to the machine. Cloud computing to process the data can be accelerated by artificial intelligence and machine learning that allow to improve the performance and reduce the computational time of the system. In the application layer, a user interface is designed to visualize data analytics and remote data monitoring in real-time. In this paper, the CONTACT Elements for IoT using the standard MQTT protocol is proposed to visualize such information through a graphical dashboard [33]. The proposed IIoT is integrated with a machine learning approach for condition monitoring of induction motor. Furthermore, the fake data due to the cyber-attack is usually the main challenge against the visualization of the real data. The cyber-attack needs to be detected in order to provide the user with real data that allows reliable decision-making and keeps the operation of the motor in a healthy condition.

### B. EXPERIMENTAL SETUP AND DATA ACQUISITION

In this section, the experimental setup and data acquisition for the induction motor is presented. Figure 3 shows the motor-driven rotary system that includes a TECO AEHF 3-phase induction, a TECO A510 series variable frequency drive, and a Chain Tail ZKB010AA magnetic particle brake. The Torque values applied on the motor were controlled using a torque meter with the type of Lorenz Dr-2477-P. The vibration signal was measured by the Wilcoxon 786A accelerometers with a sampling rate of 25,600 samples/sec. Several bearing conditions including healthy condition, outer

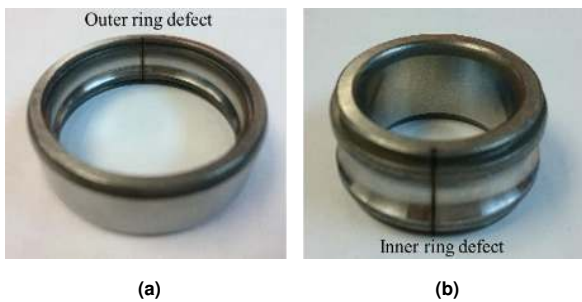


FIGURE 4. Motor defects; (a) Outer ring defect, and (b) Inner ring defect.

TABLE 1. Parameters of the experimental setup

Experiment design	Rotation speed (rpm)	Eccentricity mode	Resistance (N-m)	Failure condition
1	1350	Parallel	3.0	Normal
2	1350	Angle	0.3	Inner ring fault
3	1350	Normal	1.5	Outer ring fault
4	1800	Angle	1.5	Normal
5	1800	Normal	3.0	Inner ring fault
6	1800	Parallel	0.3	Outer ring fault

ring fault, and inner ring fault were conducted in the experiment. Additionally, fake data is also considered in this study as a cyber-attack. Figure 4 describes the outer ring and inner ring defects with 0.2 mm depth and 0.2 mm width were created by making a groove in each ring using wire electrical discharge machining.

A series of six-type experiment was conducted with three different bearing conditions, in which two different rotational speed, three electricity modes, and three resistance conditions were applied to the induction motor operations. The details of each condition are listed in Table 1. A LabVIEW program was developed to capture the vibration signals under different bearing conditions including normal condition, inner ring fault, and outer ring fault. The vibration signal was measured once the motor speed reached a steady state. Each bearing condition was recorded for a span of 20 s and it was repeated three times for each experiment condition.

The characteristic frequencies of the bearing faults can be described by Eqs. (1) and (2).

$$f_{ORF} = \frac{N_B}{2} f_R \left( 1 - \frac{D_B \cos(\theta)}{D_P} \right) \quad (1)$$

$$f_{IRF} = \frac{N_B}{2} f_R \left( 1 + \frac{D_B \cos(\theta)}{D_P} \right) \quad (2)$$

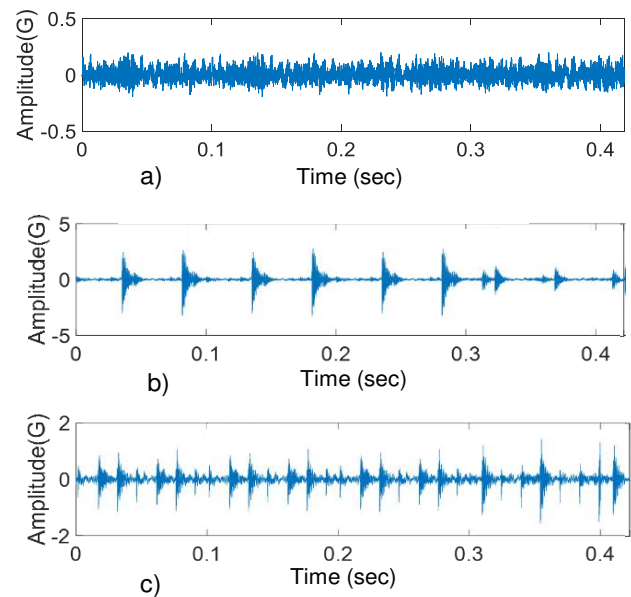


FIGURE 5. Measurement of vibration signal under different conditions (a) healthy bearing, (b) inner ring fault, and (c) outer ring fault.

where  $f_{ORF}$  and  $f_{IRF}$  are the characteristic frequencies of the outer bearing fault and inner bearing fault.  $N_B$  is the number of balls,  $D_P$  is the pitch diameter,  $D_B$  is the ball diameter,  $\theta$  is the contact angle, and  $f_R$  is the rotation speed. The 6204-T1 bearing was used in this experiment, the parameters and dimensions of the roller bearing are described in [19]. The vibration signal of each experiment was collected and labeled for each bearing condition. The time-domain spectrum of measured vibration signals under different bearing conditions is presented in Fig. 5. It shows that when the bearing failures occur, the vibration amplitudes are significantly increased in the time response compared to the response of healthy bearing. Moreover, different periods of shockwave are also observed in the cases of the inner ring and outer ring defects owing to the collision of the balls and the defect on the rings, as shown in Figs. 5 (b) and (c).

### III. MACHINE LEARNING ALGORITHMS

Recently, machine learning approaches have been applied to different engineering applications [34-37], which in turn have been widely used for bearing diagnosis of induction machines. These methods can reduce the memory computation while remaining high classification accuracy. It shows advantages comparing to linear models [38-41]. There are several different machine learning algorithms for classification problems. In this study, three effective ensemble methods machine learning techniques including decision tree (DT), random forest (RF), and extreme gradient boosting (XGBoost) algorithms are used to distinguish different bearing conditions. Those algorithms can perform very well on the respective data, particularly on the time series. They have shown effectiveness in terms of high accuracy at fast speed.

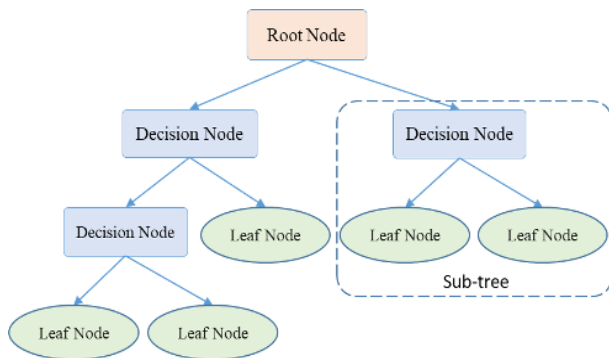


FIGURE 6. Structure of decision tree.

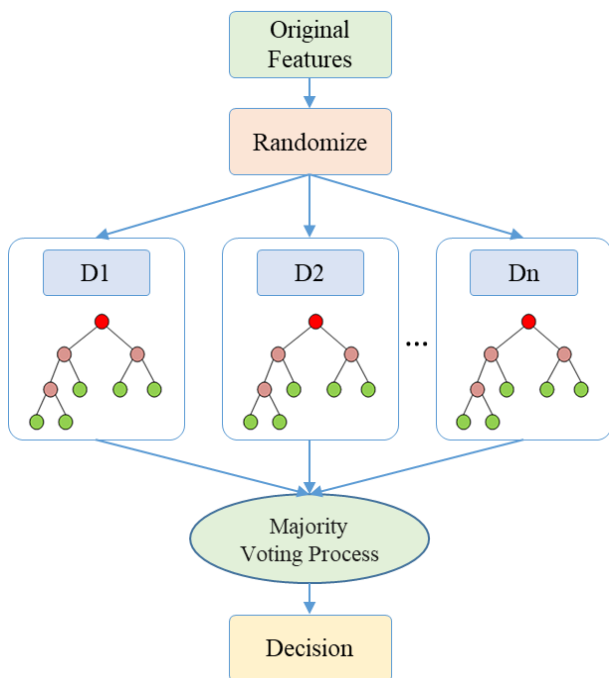


FIGURE 7. Structure of random forest for ensemble decision making.

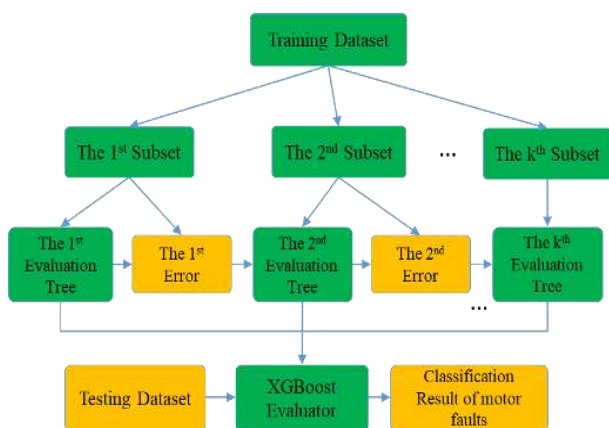


FIGURE 8. Scheme of extreme gradient boosting classifier.

Moreover, three approaches can reduce overfitting issues that make them more suitable when applying for fault detection problems compared to other machine learning algorithms such as the SVM algorithm [42].

### A. DECISION TREE

The DT method is one of the supervised learning algorithms, in which a training model is created to estimate the target from unseen samples by learning simple decision rules [43, 44]. The DT algorithm has a fast training process with low memory requirements. In the training process, to estimate the class of the given dataset, first, the values of the root attribute are compared to the real dataset attribute. The algorithm continues to compare the attribute value with the other sub-nodes in the next node and moves further. Finally, the process reaches the leaf node of the tree. Figure 6 presents the structure of the decision tree algorithm.

Suppose the training set is  $S$  and the attribute of target possesses  $n$  different values, the entropy of the training set  $S$  is described in Eq. (3).

$$Entropy(S) = \sum_{i=1}^n p_i \log_2^{p_i}, \quad (3)$$

where  $p_i$  represents the proportion of the  $i^{th}$  attribute value.

The information gain of an attribute that describes the decrease of expected entropy is defined in Eq. (4).

$$Gain(S, V) = Entropy(S) - \sum_{a \in A(V)} \frac{|S_a|}{|S|} Entropy(S_a), \quad (4)$$

where  $A(V)$  shows the range of the attribute  $A$  and  $S_a$  describes the sample set, in which the attribute  $V$  has the value of  $a$ .

In this work, a time series data set is used as a attribute or feature vector, which represents the vibration data points of a period of time. The proposed decision tree model uses the information gain, as shown in Eq. (4) to define an attribute from the attribute vector that gives maximum information about a bearing condition. The attribute with the largest information gain is selected to be the decision attribute for the node. Therefore, the level of entropy can be decreased from the root node to the leaf node.

### B. RANDOM FOREST

The random forest (RF) is also known as an effective method for fault diagnosis problems. The RF is an ensemble approach that uses tree-type classifiers. This method can enhance the performance of the model by using bagging to suppress overfitting [45]. The decisions of RF are based on the total votes of component predictors from each target. Figure 7 illustrates the structure of the random forest algorithm for ensemble decision-making.

Random forest algorithm combines many decision trees. However, an individual decision tree is now built in the forest using a random subset of attributes, and each one is trained on a random set of the training data set. Then, the prediction of the random forest model is made by average voting from every individual decision tree. Thus, the RF algorithm can reduce

overfitting in decision trees. Figure 7 illustrates the structure of the random forest algorithm for ensemble decision-making. The procedure of RF starts with the selection of random samples from a training dataset of vibration signals, then the random attribute vector consists of the limited number of attributes, which is randomly generated within the subset of data to constitute the collection. The decision tree is constructed corresponding to the elements in the collection. Each decision tree has its own decision. Finally, a majority voting process is applied to determine the optimal classification result, the prediction target with the highest score is considered to be the output.

### C. EXTREME GRADIENT BOOSTING

Extreme Gradient Boosting (XGBoost) is a scalable end-to-end tree boosting system introduced by Chen et. al. [46]. It has been widely used for fault classification problems with great performance [42, 47]. In the Gradient boosting algorithm, new models are generated to estimate the errors of previous models. Both models are combined to make the final decision, in which the stochastic gradient descent (SGD) algorithm is utilized to minimize the loss. The SGD is a very popular algorithm for optimizing an objective function using an iterative process [48]. Because the XGBoost can push the limit of computations resources for boosted tree algorithms, the number of calculations can be reduced, and the classification speed can be improved. Especially, the XGBoost classifier can suppress the overfitting problem by simplifying the objective functions. The iteration of the XGBoost algorithm starts with the first learner which is fitted to the entire data. Then, the error of the first learner will be fitted by the second learner. This process will continue the learning process and complete if a stopping condition is met. The workflow of the XGBoost classifier for motor fault diagnosis is described in Fig. 8.

## IV. RESULTS AND DISCUSSIONS

In this study, DT, RF, and XGBoost classifiers are proposed to identify four conditions of the bearing including normal, inner bearing fault, outer bearing fault, and cyber-attack case. The measured experimental data was used to generate the dataset for training and testing processes. The sample data was generated using a window size of 2000 data points for each time series, it was labeled corresponding to each bearing condition. In this study, the experimental dataset consists of 3079 samples with multiple attributes. It includes 770 samples of normal condition, 768 samples of inner ring fault, 771 samples of outer ring fault, and 770 samples of fake data. The entire dataset is split into 70% training dataset and 30% testing dataset. Note that the dataset of 770 samples is created by a randomly distributed function within the range of the motor vibration signal to represent the fake data then it is combined with the other dataset. All data were normalized using the Min-Max Scaler, which changes all features to be between 0 and 1. Then, the data was pushed into the different training models. The training and testing models were processed using a PC computer with an Intel ®

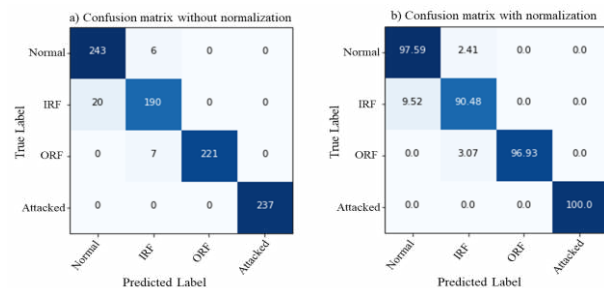


FIGURE 9. Classification result from proposed DT; (a) Confusion matrix without normalization, and (b) Confusion matrix with normalization.

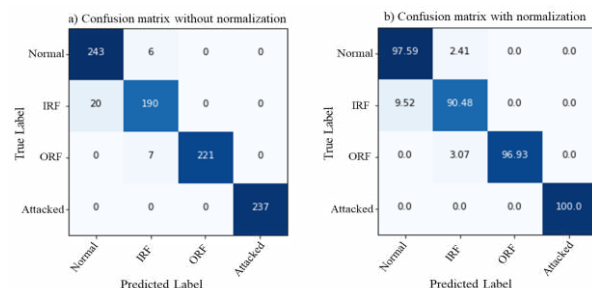


FIGURE 10. Classification result from proposed XGBoost classifier; (a) Confusion matrix without normalization, and (b) with normalization.

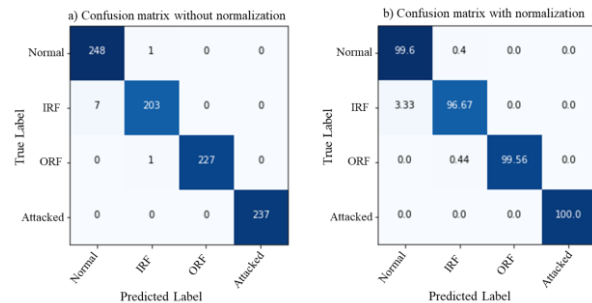


FIGURE 11. Classification result from proposed RF classifier; (a) Confusion matrix without normalization, and (b) with normalization.

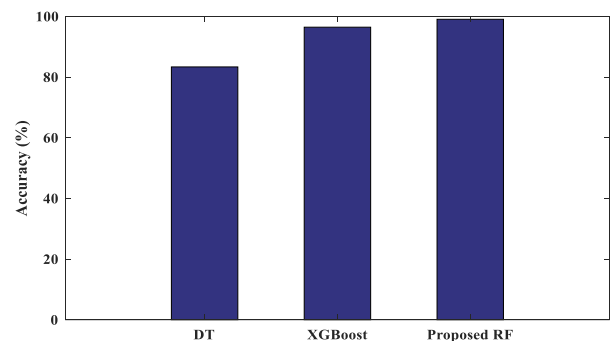


FIGURE 12. The total accuracy of each machine learning algorithm.

**TABLE 2. Accuracy and corresponding class of each proposed method**

Method \ Motor status	Normal	Outer ring faults	Inner ring faults	Attacked	Total accuracy
Decision Tree (DT)	97.59%	90.48%	96.93%	100%	83.33%
XGBoost	97.59%	90.48%	96.93%	100%	96.43%
<b>Proposed Random Forest (RF)</b>	<b>99.6%</b>	<b>96.67%</b>	<b>99.56%</b>	<b>100%</b>	<b>99.03%</b>

Core™ i7-8700 @3.20 GHz central processing unit and 6G RAM.

Generally, machine learning models have two types of parameters: model parameters and hyperparameters. While the model parameters are generated automatically by the models based on the dataset. The hyperparameters can be adjusted to setup models to improve their performance such as the learning rate parameter. Thus, the hyperparameter tuning preprocess is required to define those parameters to obtain the highest accuracy and the minimum error when constructing machine learning models. There are several optimization algorithms are used for the tuning process such as grid search and random search. The grid search is known as a great approach for hyperparameter optimization that can overcome the limitation of the manual search [49]. Therefore, the grid search algorithm is utilized to find the set of optimal hyperparameters for the RF and XGBoost models in this study. As the result of the tuning process, the hyperparameter for each model is optimized and defined. For the XGBoost model, the maximum number of iterations was optimized with “n estimators” of 400, the learning rate value is 0.1 which allows the learning speed is fast while remaining good performance of the model. The maximum depth of the tree is 4 that can control overfitting. The value of “min child weight” is 5. The learning process can be optimized using the objective function “multi: softprob”. For the RF model, the maximum number of iterations was optimized with “n estimators” of 800, an entropy criterion was selected to minimize the probability of misclassification, and the number of features to consider as looking for the best split is log2. The performances of all machine learning techniques are evaluated by the following Eq. (5).

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (5)$$

(TP = true positive; TN = true negative; FP = false positive; FN = false negative)

The classification results from different proposed machine learning approaches are shown in Figs. (9-11) and summarized in Table 2. The corresponding confusion matrices of the testing dataset demonstrate that the classifiers are very successfully identifying the bearing faults and even cyber-attack presented by fake data. Moreover, Fig. 12 shows the

total accuracy of each machine learning algorithm by bar chart as a clarified way for comparison. It shows that the proposed RF algorithm can provide an excellent classification accuracy of 99.03%, in which the false positive and the false negative numbers are very small. The DT approach has the worst performance with 83.33% classification accuracy. Whereas, the XGBoost algorithm can achieve higher classification accuracy, approximately 96.43%.

The receiver operating characteristic (ROC) curve and the area under the curve (AUC) are also carried to evaluate the performance of the three models, as shown in Fig. 13. The ROC curves present the performance of classification models with respect to classification thresholds, in which the true positive rate (TPR) and the false positive rate (FPR) are defined in Eq. (6). Whereas the AUC measures the area under the entire ROC curve. Therefore, the better classification model has a higher AUC, which can successfully distinguish between motor statuses at the AUC value of 1.0.

$$TPR = \frac{TP}{TP+FN}; \quad FPR = \frac{FP}{FP+FN} \quad (6)$$

The ROC curves the AUC values of the DT, XGboost, and RF models for distinguishing four classes (0-normal, 1-inner ring fault, 2-outer ring fault, and 3-cyber-attack) are presented in Figs. 13 (a), (b), and (c) respectively. It confirms that the proposed RF model can successfully distinguish among the classes, in which the AUC values reach 1.0 for all classes, as shown in Fig. 13 (c). The DT model shows the worst performance, which has smallest AUC values of four classes are very small and ranging from 0.8 to 0.99, as shown in Fig. 12(a). The performance of the XGBoost model in terms of AUC is between the DT and RF approaches, as illustrated in Fig. 13 (b). In addition, micro-average and macro-average ROC are used to evaluated the imbalance of data. For macro-average, models are encouraged to focus on every class correctly. Whereas, the models are highly relying on the majority classes for micro-average. The ROC and AUC results of three models show that both macro-average and micro-average AUC values are the same in each model that indicates the balance data for four classes. It is concluded that the proposed RF approach is robust and stable for the fault classification.

After training and testing, the created model of the proposed RF classifier is encrypted with the IoT architecture to



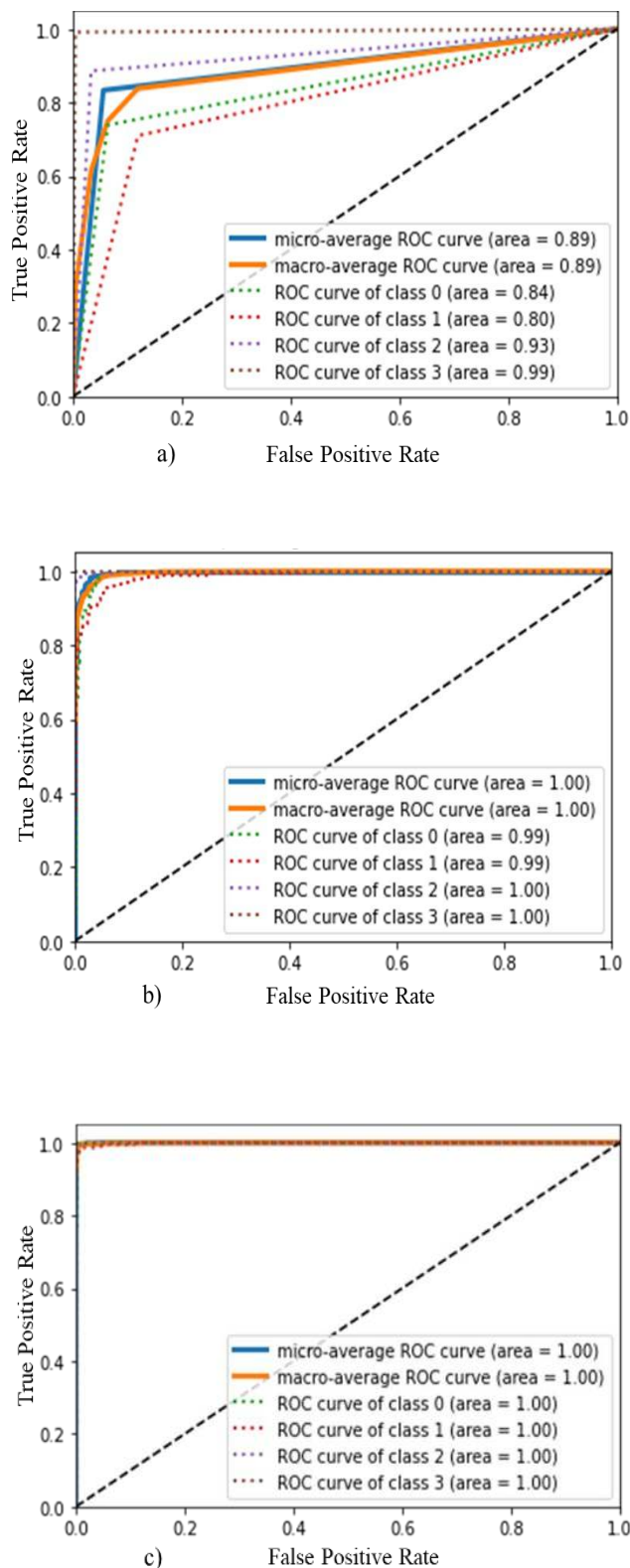


FIGURE 13. Receiver operating characteristic (ROC) curves of (a) DT classifier, (b) XGBoost classifier, and (c) RF classifier: where (0-normal, 1-inner ring fault, 2-outer ring fault, and 3-cyber-attack).

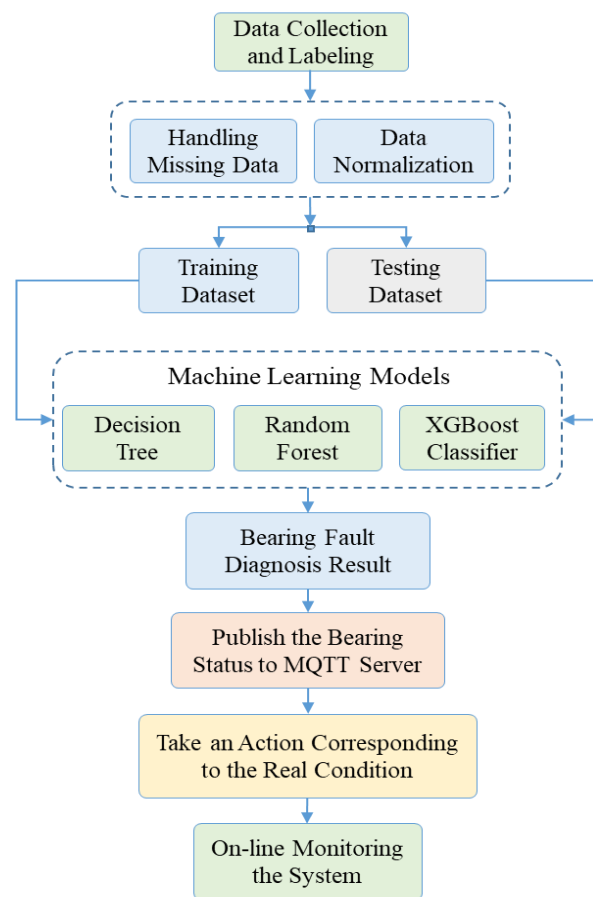


FIGURE 14. Flowchart of the proposed IoT topology with the various proposed machine learning classifiers.

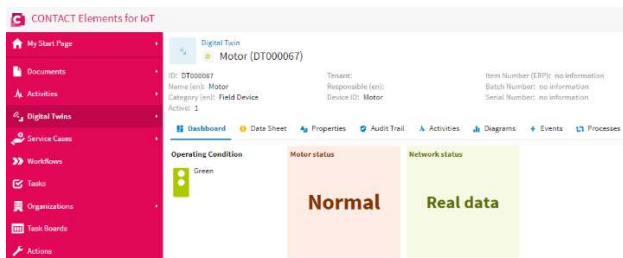
categories the online reading of the motor status and present it through an IoT dashboard as described in the following test scenarios. The flowchart in Fig. 14 describes the total operation of data acquisition, validation, and visualization. Moreover, the procedure of the cyber-attacks detection based on the proposed IoT architecture with the machine learning technique is clarified in the following pseudo-code in Algorithm 1.

#### A. SCENARIO 1: NORMAL STATE

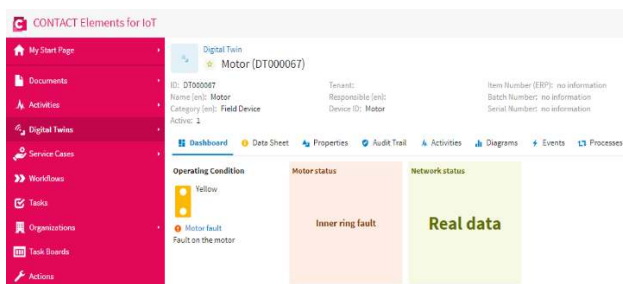
This scenario demonstrates the online monitoring for the bearing of the motor in the case of a normal state and stable network. Figure 15 presents the current status of the bearing and the reliability of the internet network on the dashboard of the CONTACT Elements for IoT. This figure shows that the motor status is normal which means the motor work well and the bearing does not have any fault. Furthermore, the light of the traffic indicator is green which means the system in a healthy state. Besides, Figure 15 shows that the internet network is stable, and the transmitted data is real. This test validates the reliability of the proposed IoT architecture and enhances the decision-making about the motor status.

**Algorithm 1.** The procedure of cyber-attack and motor faults detection

- 1: **Read** vibration signal from the motor microcontroller
- 2: **Send** signals to the IoT broker
- 3: **Input** signals to RF model
- 4: **Classify** the motor status by RF model
- 5: **Connect** to MQTT server
- 6: **if** the output of RF model = 0
- 7:     **Publish** that the motor status is ‘Normal’ and network status is ‘Real data’.
- 8: **else if** the output of RF model = 1
- 9:     **Publish** that the motor status is ‘Inner ring fault’ and network status is ‘Real data’.
- 10: **else if** the output of RF model = 2
- 11:     **Publish** that the motor status is ‘Outer ring fault’ and network status is ‘Real data’.
- 12: **else if** the output of RF model = 3
- 13:     **Publish** that the motor status is ‘Fake data’ and network status is ‘Cyber-attacks’.
- 14: **End**



**FIGURE 15.** Motor status on the IoT dashboard in the case of the normal state.

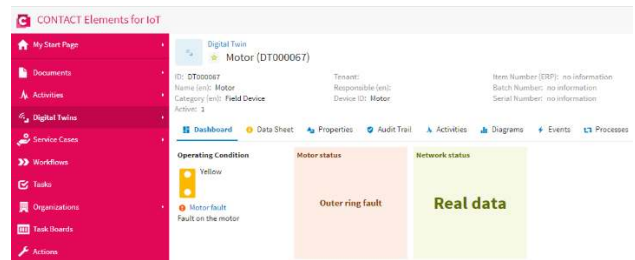


**FIGURE 16.** Motor status on the IoT dashboard in the case of inner ring fault.

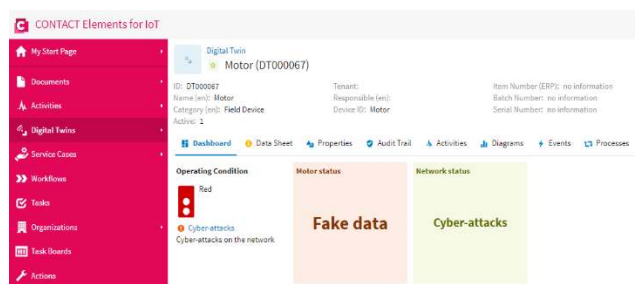
**B. SCENARIO 2: INNER RING FAULT**

This scenario is created to confirm the capability of the proposed machine learning method and the IoT architecture to detect the inner ring fault in the bearing of the motor. Figure 16 shows the bearing status in the case of inner ring fault and stable network. The IoT dashboard clears that the bearing has an inner ring fault and the transmitted data via IoT is real. Besides, the traffic light changed to a yellow color that means the motor has a fault. This test affirms the ability of the proposed machine learning method and the IoT architecture to

recognize the inner ring fault of the motor bearing without any error and clear visualization. Besides, the proposed techniques check the transmitted data against any cyber-attacks that increase the reliability of the IoT topology.



**FIGURE 17.** Motor status on the IoT dashboard in the case of outer ring fault.



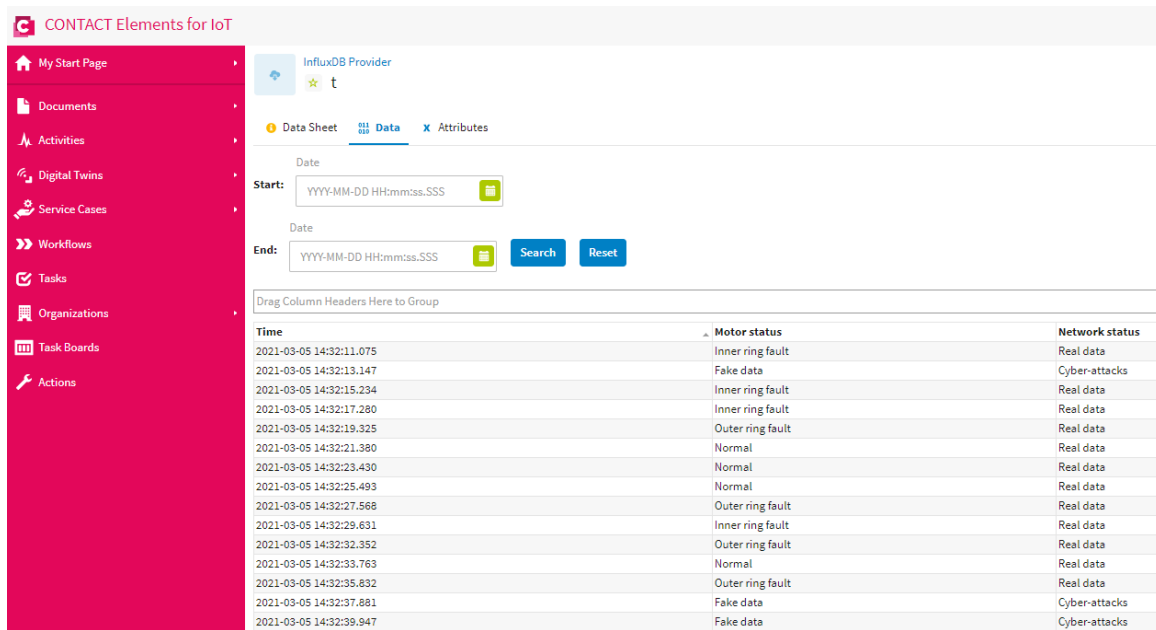
**FIGURE 18.** Network status on the IoT dashboard in the case of cyber-attacks detection.

**C. SCENARIO 3: OUTER RING FAULT**

The outer ring fault of the motor bearing is created in this scenario to demonstrates the performance of the proposed machine learning method with IoT architecture to detect all fault classes in the motor bearing. Figure 17 presents the motor bearing and the network states in the case of the outer ring fault and safe data transmission. It is clear in the IoT dashboard that the proposed machine learning technique can recognize the motor fault of the outer ring. Furthermore, the traffic light changed from the green light of the normal state to the yellow light to inform the user about the abnormal state due to the outer ring fault of the motor bearing. Besides, the dashboard shows that the network is stable, and the transmitted data is real.

**D. SCENARIO 4: CYBER-ATTACKS DETECTION**

Cyber-attacks are the main challenge against the implementation of the IoT system. So, this scenario is carried out to ensure the effectiveness of the proposed machine to recognize the cyber-attacks on the internet network. Figure 18 presents the transmitted data status in case of cyber-attacks on the internet network. It is clear from this figure that the proposed machine learning technique can detect cyber-attacks on internet networks and the proposed IoT architecture clears that the transmitted data about the motor status is fake data. Furthermore, the traffic light changed from the green light of



**FIGURE 19.** The latency/overhead of the proposed fault and cyberattack detection process based on the proposed IoT platform.

the normal state to the red light to inform the user about the cyber-attacks on the network that enhance the decision-making to maintain and stabilize the internet network. This test confirms that the proposed machine learning with the IoT architecture can provide reliable monitoring for the motor status.

The results from the above scenarios can be concluded as follows;

- The proposed IoT architecture can visualize the normal status of the motor and the internet network status clearly on the dashboard of the IoT platform as presented in scenario 1 effectively without any errors.
- The inner ring fault of the motor can be detected effectively by the proposed IoT topology based on the machine learning technique as cleared in scenario 2. In addition, the suggested IoT platform performs an alarm and converted the color of the traffic light from the green color to the yellow color to remind the user about the inner ring fault of the motor.
- The third test case shows the outer ring fault of the motor and emphasizes the superiority of the suggested IoT architecture based on the machine learning technique to recognize and visualize the outer ring fault of the motor effectively.
- The final scenario clears the superiority of the introduced machine learning technique to recognize the cyber-attacks on the network. Furthermore, the suggested IoT platform presents the status of transmitted data through the internet network if it is real or fake to enhance the decision-making about the motor status. Besides, the IoT platform converted the traffic light to red color in order to remind the user to take care of the network against cyber-attacks issue.

- The latency/overhead of the proposed fault and cyberattack detection process is analyzed by the proposed IoT platform and presented in Fig. 19.

## V. CONCLUSIONS

This paper introduces a new IoT architecture for online monitoring of the faults of the induction motor instead of the traditional methods. The proposed IoT architecture is developed based on effective machine learning techniques to recognize the fault classes of the motor. Besides, the cyber-attacks issue is taken into the account and the attack can be detected and suppressed by the proposed IoT topology. Different experimental testing is performed to confirm the effectiveness of the proposed IoT architecture based on machine learning. The results emphasizes the superiority of the proposed IoT architecture to recognize motor faults and cyber-attacks with high accuracy. Furthermore, the results are visualized on CONTACT Elements for IoT platform clearly that enhance the decision making about the motor statues. That allows the system to run for a long time in healthy conditions and improve the industrial environment that represents the main target of Industry 4.0. In addition, the proposed IoT architecture provides a promising solution to apply to different machines for future work.

## ACKNOWLEDGMENTS

The authors acknowledge the CONTACT Elements for IoT platform for supporting this work that applied in online monitoring of fault diagnosis in induction motors.

## REFERENCES

- [1] A. Singh, B. Grant, R. DeFour, C. Sharma, and S. Bahadoorsingh, "A review of induction motor fault modeling," *Electric Power Systems Research*, vol. 133, pp. 191-197, 2016/04/01/ 2016.
- [2] Y. Liu and A. M. Bazzi, "A review and comparison of fault detection and diagnosis methods for squirrel-cage induction motors: State of the art," *ISA Transactions*, vol. 70, pp. 400-409, 2017/09/01/ 2017.
- [3] Choudhary, A., Goyal, D., Shimi, S.L. et al. Condition Monitoring and Fault Diagnosis of Induction Motors: A Review. *Arch Computat Methods Eng* 26, 1221–1238 (2019). <https://doi.org/10.1007/s11831-018-9286-z>.
- [4] Blodt, Martin & Granjon, Pierre & Raison, Bertrand & Regnier, Jeremi. (2010). Mechanical Fault Detection in Induction Motor Drives through Stator Current Monitoring - Theory and Application Examples. 10.5772/9072.
- [5] F. Cipollini, L. Oneto, A. Coraddu, and S. Savio, "Unsupervised Deep Learning for Induction Motor Bearings Monitoring," *Data-Enabled Discovery and Applications*, vol. 3, no. 1, p. 1, 2019/01/04 2019.
- [6] F. Immovilli, C. Bianchini, M. Cocconcelli, A. Bellini, and R. Rubini, "Bearing Fault Model for Induction Motor With Externally Induced Vibration," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3408-3418, 2013.
- [7] J. J. Saucedo-Dorantes, M. Delgado-Prieto, J. A. Ortega-Redondo, R. A. Osornio-Rios, and R. d. J. Romero-Troncoso, "Multiple-Fault Detection Methodology Based on Vibration and Current Analysis Applied to Bearings in Induction Motors and Gearboxes on the Kinematic Chain," *Shock and Vibration*, vol. 2016, p. 5467643, 2016/03/13 2016.
- [8] M. Iorgulescu and R. Beloiu, "Vibration and Current Monitoring for Fault's Diagnosis of Induction Motors," 2008.
- [9] M.-K. Liu, M.-Q. Tran, and P.-Y. Weng, "Fusion of Vibration and Current Signatures for the Fault Diagnosis of Induction Machines," *Shock and Vibration*, vol. 2019, p. 7176482, 2019/09/19 2019.
- [10] ISO-10816, "INTERNATIONAL STANDARD TIME," (in eng), *Science*, vol. 1, no. 6, pp. 159-60, Mar 16 1883.
- [11] M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, and H. Henao, "Bearing Fault Detection by a Novel Condition-Monitoring Scheme Based on Statistical-Time Features and Neural Networks," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 8, pp. 3398-3407, 2013.
- [12] D. Wu *et al.*, "An Automatic Bearing Fault Diagnosis Method Based on Characteristics Frequency Ratio," *Sensors*, vol. 20, no. 5, p. 1519, 2020.
- [13] P. Singh and S. Harsha, "Statistical and frequency analysis of vibrations signals of roller bearings using empirical mode decomposition," *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics*, vol. 233, no. 4, pp. 856-870, 2019.
- [14] H.-C. Lin and Y.-C. Ye, "Reviews of bearing vibration measurement using fast Fourier transform and enhanced fast Fourier transform algorithms," *Advances in Mechanical Engineering*, vol. 11, no. 1, p. 1687814018816751, 2019.
- [15] Q. Hu, Z. He, Z. Zhang, and Y. Zi, "Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble," *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 688-705, 2007/02/01/ 2007.
- [16] Z. K. Peng, F. L. Chu, and P. W. Tse, "Singularity analysis of the vibration signals by means of wavelet modulus maximal method," *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 780-794, 2007/02/01/ 2007.
- [17] Z. German-Sallo and H. S. Grif, "Hilbert-Huang Transform in Fault Detection," *Procedia Manufacturing*, vol. 32, pp. 591-595, 2019/01/01/ 2019.
- [18] P. K. Kankar, S. C. Sharma, and S. P. Harsha, "Rolling element bearing fault diagnosis using wavelet transform," *Neurocomputing*, vol. 74, no. 10, pp. 1638-1645, 2011/05/01/ 2011.
- [19] M.-K. Liu and P.-Y. Weng, "Fault Diagnosis of Ball Bearing Elements: A Generic Procedure based on Time-Frequency Analysis," *Measurement Science Review*, vol. 19, pp. 185-194, 08/01 2019.
- [20] I. Cruz-Vega, J. Rangel-Magdaleno, J. Ramirez-Cortes, and H. Peregrina-Barreto, "Automatic progressive damage detection of rotor bar in induction motor using vibration analysis and multiple classifiers," *Journal of Mechanical Science and Technology*, vol. 31, no. 6, pp. 2651-2662, 2017/06/01 2017.
- [21] B.-S. Yang, X. Di, and T. Han, "Random forests classifier for machine fault diagnosis," *Journal of Mechanical Science and Technology*, vol. 22, no. 9, pp. 1716-1725, 2008/09/01 2008.
- [22] J. Xie, Z. Li, Z. Zhou, and S. Liu, "A Novel Bearing Fault Classification Method Based on XGBoost: The Fusion of Deep Learning-Based Features and Empirical Features," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-9, 2021.
- [23] R. Nishat Toma and J.-M. Kim, "Bearing Fault Classification of Induction Motors Using Discrete Wavelet Transform and Ensemble Machine Learning Algorithms," *Applied Sciences*, vol. 10, no. 15, p. 5251, 2020.
- [24] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.
- [25] Y. Ma *et al.*, "A new fault diagnosis method based on convolutional neural network and compressive sensing," *Journal of Mechanical Science and Technology*, vol. 33, no. 11, pp. 5177-5188, 2019/11/01 2019.
- [26] O. Janssens *et al.*, "Convolutional Neural Network Based Fault Detection for Rotating Machinery," *Journal of Sound and Vibration*, vol. 377, pp. 331-345, 2016/09/01/ 2016.
- [27] L. Xiao, Y. Li, X. Huang and X. Du, "Cloud-Based Malware Detection Game for Mobile Devices with Offloading," *IEEE Transactions on Mobile Computing*, vol. 16, no. 10, pp. 2742-2750, Oct. 2017, doi: 10.1109/TMC.2017.2687918.
- [28] F. Civerchia, S. Bocchino, C. Salvadori, E. Rossi, L. Maggiani, and M. Petracca, "Industrial Internet of Things monitoring solution for advanced predictive maintenance applications," *Journal of Industrial Information Integration*, vol. 7, pp. 4-12, 2017/09/01/ 2017.
- [29] R. Y. Zhong, L. Wang, and X. Xu, "An IoT-enabled Real-time Machine Status Monitoring Approach for Cloud Manufacturing," *Procedia CIRP*, vol. 63, pp. 709-714, 2017/01/01/ 2017.
- [30] Q. Cao, F. Giustozzi, C. Zanni-Merk, F. de Bertrand de Beuvron, and C. Reich, "Smart Condition Monitoring for Industry 4.0 Manufacturing Processes: An Ontology-Based Approach," *Cybernetics and Systems*, vol. 50, no. 2, pp. 82-96, 2019/02/17 2019.
- [31] R. Y. Zhong, X. Xu, E. Klotz, and S. T. Newman, "Intelligent Manufacturing in the Context of Industry 4.0: A Review," *Engineering*, vol. 3, no. 5, pp. 616-630, 2017/10/01/ 2017.
- [32] S. Phuyal, D. Bista, and R. Bista, "Challenges, Opportunities and Future Directions of Smart Manufacturing: A State of Art

- Review," *Sustainable Futures*, vol. 2, p. 100023, 2020/01/01/2020.
- [33] (Dec. 22, 2020). IoT platform for digital business models | CONTACT Software, <https://www.contact-software.com/en/products/iot-platform-for-digital-business-models/?fbclid=IwAR0oYDd4qHpCd0BEZaGrLHEAQGYoQ2BhBmDzbF35-cyM6QrNHakziWDC8yo>.
- [34] M. Elsis, M.-Q. Tran, K. Mahmoud, M. Lehtonen, and M. M. F. Darwish, "Deep Learning-Based Industry 4.0 and Internet of Things towards Effective Energy Management for Smart Buildings," *Sensors*, vol. 21, no. 4, p. 1038, Feb. 2021.
- [35] M. Elsis, M. -Q. Tran, K. Mahmoud, D. -E. A. Mansour, M. Lehtonen and M. M. F. Darwish, "Towards Secured Online Monitoring for Digitalized GIS Against Cyber-Attacks Based on IoT and Machine Learning," in *IEEE Access*, doi: 10.1109/ACCESS.2021.3083499.
- [36] M. Elsis, K. Mahmoud, M. Lehtonen, and M. M. F. Darwish, "Reliable Industry 4.0 Based on Machine Learning and IoT for Analyzing, Monitoring, and Securing Smart Meters," *Sensors*, vol. 21, no. 2, p. 487, Jan. 2021.
- [37] M. Elsis, M. -Q. Tran, K. Mahmoud, M. Lehtonen and M. M. F. Darwish, "Robust Design of ANFIS-Based Blade Pitch Controller for Wind Energy Conversion Systems Against Wind Speed Fluctuations," in *IEEE Access*, vol. 9, pp. 37894-37904, 2021, doi: 10.1109/ACCESS.2021.3063053.
- [38] J. G. Ferreira and A. Warzecha, "An application of machine learning approach to fault detection of a synchronous machine," in *2017 International Symposium on Electrical Machines (SME)*, 2017, pp. 1-6.
- [39] M. N. Ali, K. Mahmoud, M. Lehtonen and M. M. F. Darwish, "An Efficient Fuzzy-Logic Based Variable-Step Incremental Conductance MPPT Method for Grid-Connected PV Systems," in *IEEE Access*, vol. 9, pp. 26420-26430, 2021, doi: 10.1109/ACCESS.2021.3058052.
- [40] M. Elsis, K. Mahmoud, M. Lehtonen and M. M. F. Darwish, "An Improved Neural Network Algorithm to Efficiently Track Various Trajectories of Robot Manipulator Arms," in *IEEE Access*, vol. 9, pp. 11911-11920, 2021, doi: 10.1109/ACCESS.2021.3051807.
- [41] M. Elsis, K. Mahmoud, M. Lehtonen and M. M. F. Darwish, "Effective Nonlinear Model Predictive Control Scheme Tuned by Improved NN for Robotic Manipulators," in *IEEE Access*, vol. 9, pp. 64278-64290, 2021, doi: 10.1109/ACCESS.2021.3075581.
- [42] D. Zhang, L. Qian, B. Mao, C. Huang, B. Huang, and Y. Si, "A Data-Driven Design for Fault Detection of Wind Turbines Using Random Forests and XGboost," *IEEE Access*, vol. 6, pp. 21020-21031, 2018.
- [43] Y. Zhao, L. Yang, B. Lehman, J. d. Palma, J. Mosesian, and R. Lyons, "Decision tree-based fault detection and classification in solar photovoltaic arrays," in *2012 Twenty-Seventh Annual IEEE Applied Power Electronics Conference and Exposition (APEC)*, 2012, pp. 93-99.
- [44] R. Benkercha and S. Moulahoum, "Fault detection and diagnosis based on C4.5 decision tree algorithm for grid connected PV system," *Solar Energy*, vol. 173, pp. 610-634, 2018/10/01/2018.
- [45] C. Aldrich and L. Auret, "Fault detection and diagnosis with random forest feature extraction and variable importance methods," *IFAC Proceedings Volumes*, vol. 43, no. 9, pp. 79-86, 2010/01/01/2010.
- [46] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," presented at the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA, 2016. Available: <https://doi.org/10.1145/2939672.2939785>
- [47] Z. Wu, X. Wang, and B. Jiang, "Fault Diagnosis for Wind Turbines Based on ReliefF and eXtreme Gradient Boosting," *Applied Sciences*, vol. 10, no. 9, 2020.
- [48] Friedman, Jerome. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*. 29. 1189-1232.
- [49] Yu, T., & Zhu, H. (2020). Hyper-parameter optimization: A review of algorithms and applications. arXiv preprint arXiv:2003.05689.