IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 10 August 2023, accepted 21 September 2023, date of publication 25 September 2023,
date of current version 17 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3318866

RESEARCH ARTICLE

A Novel Federated Edge Learning Approach for Detecting Cyberattacks in IoT Infrastructures

SIDRA ABBAS[1], ABDULLAH AL HEJAILI[2], GABRIEL AVELINO SAMPEDRO[3,4], (Member, IEEE), MIDETH ABISADO[5], AHMAD S. ALMADHOR[6], (Member, IEEE), TARIQ SHAHZAD[7], AND KHMAIES OUAHADA[7]

[1]Department of Computer Science, COMSATS University Islamabad, Islamabad 45550, Pakistan
[2]Faculty of Computers and Information Technology, Computer Science Department, University of Tabuk, Tabuk 71491, Saudi Arabia
[3]Faculty of Information and Communication Studies, University of the Philippines Open University, Los Baños 4031, Philippines
[4]Center for Computational Imaging and Visual Innovations, De La Salle University, Manila 1004, Philippines
[5]College of Computing and Information Technologies, National University, Manila 1008, Philippines
[6]Department of Computer Engineering and Networks, College of Computer and Information Sciences, Jouf University, Sakaka 72388, Saudi Arabia
[7]Department of Electrical and Electronic Engineering Science, University of Johannesburg, Johannesburg 2006, South Africa

Corresponding author: Khmaies Ouahada (kouahada@uj.ac.za)

This work was supported by the University of Johannesburg.

ABSTRACT The advancement of the communications system has resulted in the rise of the Internet of Things (IoT), which has increased the importance of cybersecurity research. IoT, which incorporates a range of devices into networks to offer complex and intelligent services, must maintain user privacy and deal with attacks such as spoofing, denial of service (DoS), jamming, and eavesdropping. Attacks change with time, and new ones develop every day. Numerous researchers look into IoT system attack models and evaluate machine, deep, and federated learning-based IoT security approaches. However, existing methods do not produce reliable and encouraging performance. Therefore, this study proposes a novel approach for leveraging federated learning to identify large attacks on IoT devices using the novel CIC_IoT 2023 dataset. The approach uses a federated deep neural network to achieve precise categorization. Before model training, the data was preprocessed using various data preparation techniques to guarantee the creation of a trustworthy dataset for categorization. The suggested approach involves feature normalization, data balancing, and model prediction utilizing federated learning. The experimental findings show that the proposed approach attained an exceptional accuracy of 99.00%, endorsing it for attack detection.

INDEX TERMS Internet of Things (IoT), networks attacks, privacy, preservational deep learning, federated learning.

I. INTRODUCTION

The Internet of Things (IoT) has acquired popularity and is already impacting the fortune. The IoT intends to convert everyday life by creating billions of smart devices, roughly 75 billion IoT gadgets by 2025, to carry out routine chores. Various industries, including healthcare, agriculture, transportation, and manufacturing, are starting to rely heavily on the IoT [1]. The IoT is spreading like wildfire. During the past few years, the number of connections to various items in industry, residences, and transportation has all expanded quickly. IOT-enabled devices come in all shapes and sizes, using different appliances with microcontrollers to measure temperature. These devices provide data for analysis to the cloud server via internet communication [2]. IoT and further communication technologies have significantly enhanced our understanding of our surroundings. IoT technologies can collect and interpret data about the surrounding environment and enhance life quality. This situation makes it simpler for people and objects to communicate, promoting the growth of smart cities [3].

The IoT is a complex, networked procedure because of the huge attack surface; it is challenging to assemble the security needs of an IoT system. IoT deployment has become an

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

interconnected process due to the IoT's extensive use. When establishing an IoT system, many factors must be considered, including security, energy capability, analytics techniques, and interoperability with different software programs. On the other side, IoT devices frequently function without mortal operators. Hence, a stealer might physically approach these devices. Eavesdropping on wireless networks utilized by IoT devices, which are frequently linked by a communication medium, allows intruders access to sensitive information [4]. Numerous studies on the IoT network have demonstrated that privacy and data security are the two most crucial factors to consider. Devices will also deliver a gateway into the broadcasted network with a feeble point of security against invading devices [5], [6].

A. RESEARCH MOTIVATION

Artificial intelligence is an emerging technology nowadays; researchers have used machine and deep learning methodologies for prediction. Researchers used different classification approaches to detect attacks or malware in IoT networks, such as Long Short-Term Memory (LSTM) [7], ShieldRNN [8], Deep Neural Networks (DNN) [9], Bridge Structure Health Monitoring (BSHM) [10] and Radial basis Function Network (RBFN) [11]. These techniques are limited in providing better security and privacy to clients' data. IoT-based systems require huge information transfers to improve the structure and user experience. Typically, sensors in the system are used to gather these extremely sensitive user data. These smart devices depend on data supplied or obtained through the network as IoT-based systems develop. Data collection and transmission from the user to the client puts it at risk of attack. The system's security becomes compromised if these data are changed [12]. In such a scenario, researchers proposed a methodology that transmits the outcomes to the main server rather than processing every obtained user's data in a particular client. The operation is finished, and data security is maintained adequately when data is encrypted and held solely on the client side, and the outcome is sent to the server [13].

B. RESEARCH CONTRIBUTION

This study proposes an approach to detect attacks on IoT devices more precisely and efficiently. The research's main contributions and the steps that are carried out are listed below.

- Present a federated learning approach to predict an attack based on enhancing IoT device security using the CIC_IOT dataset.
- This study employs a two-client, one-server architecture in which each client individually trains its model utilizing its data before updating it on the server. The server compiles the modifications and provides users with the updated global model. This architecture makes distributed training possible while the server's processing load is reduced and data privacy is maintained.
- Data preparation (data balancing and feature normalization) is used anomalously for classifying after obtaining the data. A deep neural network classifier forecasts different types of attacks to stop them and offer protection.
- The experimental results reveal that the proposed approach attained an exceptional accuracy of 99.00%, endorsing it for attack detection.

C. RESEARCH ORGANIZATION

The most recent pertinent research on IoT devices security prediction, machine learning, deep learning and federated learning approaches are presented in Section II. Dataset selection, data preprocessing (data balancing, feature normalization), federated learning technique (model architecture) and evaluation measurements are discussed in Section III. The experimental analysis and result are discussed in Section IV, in which results of the experimental server side, client 1 and the client are explained in the proposed methodology. Section V concludes and makes recommendations for additional research.

II. RELATED WORK

An overview of relevant publications and studies in the domain of machine learning, deep learning and federated learning are explained in this section.

A. MACHINE AND DEEP LEARNING TECHNIQUES

Author in [8] a practical defense against such unavoidable attacks on IoT devices. A solution comprises an IoT node sensor and a server sensor. The IoT node sensor is a simple classifier that keeps track of exit traffic. If the IoT node suspects it is part of a Distributed Denial-of-Service (DDoS) attack, it will employ the server sensor, a better precise classifier. This research proposed ShieldRNN, a unique training and forecast strategy for RNN/LSTM models, to create an accurate server detector. On the CIC-IDS2017 dataset, they corresponded with ShieldRNN with other supervised and unsupervised techniques and demonstrated their superior performance. Authors in [10] investigated a key topic inspired by useful IoT-based BSHM: extending the network's lifespan while effectively ensuring target coverage. This paper proposed an energy-efficient detector scheduling method to achieve partial CIC scope in IoT-based BSHM systems to ensure network range and lengthen the network's lifespan. The strategy integrated a promising reinforcement learning technique, Learning Automata (LA), with the CIC technique. While fulfilling network connection and partial coverage ratio, the suggested strategy completely utilizes cooperation between deployed nodes and alternately schedules the wake/sleep level of nodes.

Authors in [14] validated the IoTDevID technique by putting its main features and aggregation process to the test on the CIC-IoT2022 dataset. The dataset has several enhancements over past datasets, including more devices, numerous instances of the same device, data on both IP and non-IP devices, information on typical (benevolent)

usage, and various usage profiles, including active and idle states. With a 92.50 F1 score for 31 IP-only device classes, IoTDevID's results are strong and consistent with those from earlier datasets. The IoTDevID aggregation approach enhanced model performance in every instance. Even with significantly less data, we can still acquire a 78.80 F1 score for 40 different device classes for non-IP devices, demonstrating that data volume also affects model performance. Authors in [15] paper presented a DoS prediction technique based on machine learning (ML). Based on instances obtained by the sFlow protocol straight from network appliances, the software utilizes the Random Forest Tree classifier to categorize network traffic. Four recent benchmark datasets are utilized in the tests. Employing an instance rate of 20% of network traffic, the outcomes show an online prediction ratio of attacks above 96% with high performance and a low false alarm rate.

Authors in [16] proposed an innovative approach for identifying fake data injection attacks during state estimation. The state vectors' spatial and secular data correlations may differ from those under typical operating settings when malicious data are inserted into them. The suggested mechanism analyses temporally successive anticipated system conditions employing wavelet convert and deep neural network techniques to catch such inconsistency successfully. With thorough case studies on IEEE 118- and 300-bus power systems, they evaluate the performance of the presented approach. The findings show that the method can detect attacks with a fair degree of accuracy. Authors in [17] thoroughly analyze feature groups' significance and detective capability for identifying network attacks. Three feature selection models–chi-square, information gain, and correlation, have been used to find and organize data features. To assess the effectiveness of two ML classifiers for detecting attacks, deep feed-forward and random forest are fed the attributes. In their unique flow format, three datasets, UNSW-NB15, CSE-CIC-IDS2018, and ToN-IoT, are considered for practical evaluation. The corresponding NetFlow format variants NF-UNSW-NB15, NF-CSE-CIC-IDS2018, and NF-ToN-IoT are also considered. The findings indicate that while accuracy initially rises quickly with adding characteristics, it converges to its maximum.

B. FEDERATED LEARNING

Authors in [18] propose using FELIDS, a federated learning-based intrusion detection system, to safeguard agricultural IoT infrastructures. The FELIDS system uses three DL classifiers, DNN, Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN), to guard against Agricultural IoTs threats. They examine the effectiveness of the suggested IDS using data from the CSE-CIC-IDS2018, MQTTset, and InSDN. The outcomes show that the FELIDS system obtains the maximum accuracy in identifying assaults and surpasses the traditional/centralized arrangements of machine learning (non-federated learning) in safeguarding the privacy of IoT device data.

Authors in [13] thoroughly assess the differential privacy techniques used when an FL-enabled IDS is trained for industrial IoT. They examine the accuracy gained while considering various privacy criteria and aggregation functions, notably FedAvg and the recently proposed Fed+, in contrast to prior approaches that dealt with non-independent and identically distributed data over the latest ToN_IoT dataset. Analysis shows that employing Fed+ in the environment yields comparable outcomes even when noise is added during the federated training phase. Authors in [19] provided a brand-new FL-empowered semi-supervised active learning (FL-SSAL) security orchestration model for the Label-at-Client problem, in which clients also have some unlabeled samples but a smaller amount of labeled data. Entropy-based active learning uses unlabeled data semi-supervised while choosing the most informative samples for data annotation. Their experimental assessments on the confidential, non-independent, and identically dispersed (non-IID) dataset show that FL-SSAL outperforms baseline systems with less labeled data regarding intrusion detection accuracy while having a lower communication overhead.

III. PROPOSED APPROACH

This section demonstrates the methodology discussed in depth based on dataset selection, dataset pre-preparation, data balancing, feature normalizing, algorithms and performance evaluation metrics. Figure 1 depicts the overall process used in this research for large-scale attack detection in IoT devices. We apply a data preparation technique that balances the data using SMOTE. Data normalization is done using the MinMaxScaler technique. To improve results, this study used the federated learning technique to boost a deep learning model's performance in a distributed setting. A deep neural network was used in the study to classify and identify attacks on IoT devices from the CIC_IoT dataset.

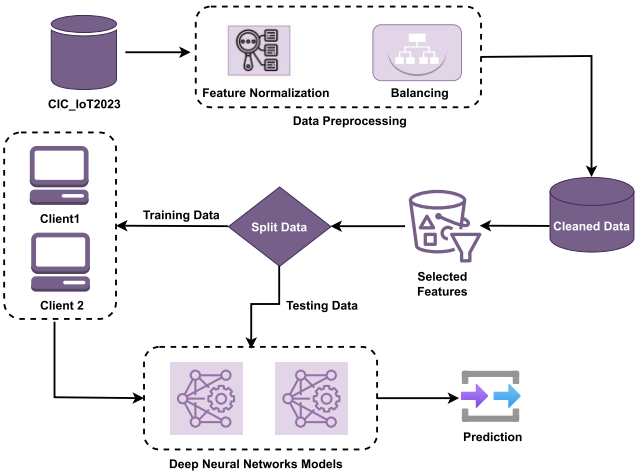


FIGURE 1. Proposed approach for IoT attack detection.

The proposed attack detection on the IoT device approach algorithm involves various steps, as illustrated in Algorithm 1. The initial step in preparing the dataset is to

address the class imbalance, after which data balancing is carried out. The minority population might be oversampled, or the majority population can be undersampled to achieve this. The normalization of features uses the min-max scaler. Returning the prepared, normalized dataset together with the pertinent labels. The DNN model is then trained after that. An input layer with an input shape is part of the model architecture (46). Two hidden layers, each with 64 and 32 units, are presented. A fully linked layer with one unit and sigmoid activation is used as the output layer. The binary cross-entropy loss function and Adam optimizer are used to train the model. The model is used to generate predictions on test data after training. As a prediction, the Y_predicted labels are given back. The main method then initializes a data frame with the CIC_IoT data. The dataset is created using the ''Prepare Dataset'' function, and the DNN model is trained using the ''TrainDNN'' function. The ''ModelPrediction'' function is used on test data to forecast models. The equation displayed in algorithm 1 determines the model's accuracy. The final outputs of the algorithm are the predicted labels Y_predicted and the accuracy.

Algorithm 1 Proposed Federated Learning Approch

Input: D_s (CIC_IoT dataset)
Output: Attacks prediction on IoT devices
Function Prepare dataset D_p (CIC_IoT Data)
Perform data balancing D_b
Feature normalization: $Y_{norm} = Y_i$-Y_{min}/Y_{max}-Y_{min}
Return Y_{norm}, y
Function TrainDNN (Y_{norm}, y)
  Start Model Training M_t
  Set Input Shape I_s (46,)
  Add 64-unit and 32-unit hidden dense layers
  Add dropouts hidden layers
  Add 1 unit, fully connected layer, sigmoid activation
  Use Binary Cross-entropy, loss and Adam Function
Return TrainDNN
Function ModelPrediction (TrainDNN, test_data (T_d))
  Start Model Prediction
  $Y_{Prediction}$ = TrainDNN.predict(T_d)
Return $Y_{Prediction}$
Working Main
  $D_s \leftarrow$ Data initialization
  Y_{norm}, y $\leftarrow$ Dataset preparation (D_s)
  $Model \leftarrow$ TrainDNN (Y_{norm}, y)
  $Y_{Prediction} \leftarrow$ ModelPrediction ($Model$, T_d)
  Performance computing using Accuracy
Return Performance, $Y_{Prediction}$

A. DATASET SELECTION AND PREPROCESSING

In this research, we use CICIoT2023, an IoT attack dataset based on an extensive topology composed of several real IoT devices acting as either attackers or victims. The study's author demonstrated how to reproduce 33 attacks against IoT devices broken down into seven classes, executed, documented, and collected data on them [20]. This study uses two preprocessing steps to organize the data before training the model. These two preprocessing steps are explained below:

Data Balancing: Data resampling technique Synthetic Minority Oversampling Technique (SMOTE) is employed in this study to give more representative samples [21]. To create synthetic samples, random observations from the minority class are chosen, and these observations are then subtly altered to produce new data points. This upsampling technique is created to address the class imbalance in the initial dataset and increase the effectiveness of the machine learning model. By generating additional data from the minority class, we ensure that the model is neutral and can accurately forecast both classes.

Feature Normalization: is a technique to rescale the dataset's features between 0 and 1, helping to standardize and make them comparable [22]. The dataset's features are converted into a specified range, in this case, spanning from 0 to 1, using the MinM axScaler technique. Using this method, the original value of each feature is subtracted from its minimum value, which is then divided by the difference between the feature's maximum and minimum values.

$$Y_{norm} = \frac{Y_j - Y_{min}}{Y_{max} - Y_{min}} \tag{1}$$

Equation 1 expresses the mathematical form of normalization. Here, Y stands for the feature's actual value, Ynorm for its normalized value, Ymin for the feature's smallest value in the dataset, and Ymax for its highest value.

B. FEDERATED LEARNING TECHNIQUE

Federated learning permits several clients to train a single model together without exchanging local data. In this scenario, data features are used by two clients to build a desired deep-learning model. The initial process involves taking client data and running it through data preprocessing steps to pull out useful features. The clients then use the deep learning model already provided to them. This model consists of a sigmoid-activated dense output layer, three further dense layers, an input and two dropout layers. Kernel regularizer with 0.1 value, relu activation function is used as hyperparameter in hidden layers. When clients transfer their features to the central server, training can start. The server then gathers the feature and modifies the model weights following the data received. The buyers are then given the latest version of the model. Each user keeps using their information to improve the updated model. This aggregation and model updating process is carried out iteratively to reach convergence. The model is trained using the accuracy measurement, the Adam optimizer, and the binary categorical cross-entropy loss function. A batch size of 32 spanning 5 epochs is used for the training procedure. Using a validation split of 10%, the model's performance is assessed during and after training. Clients can alter model training using this

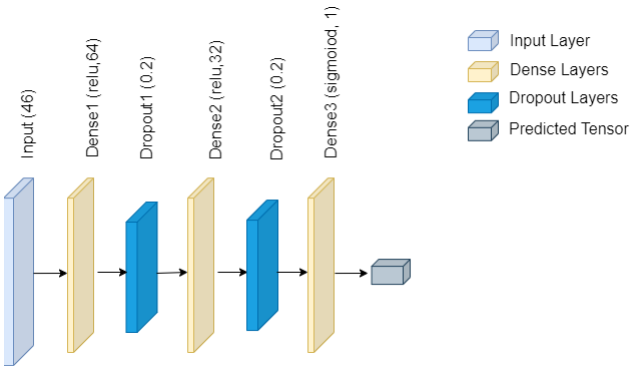


FIGURE 2. Deep neural network architecture.

federated learning technique while maintaining the privacy of their local data. The subsequent attack on IoT devices is predicted using the created model.

Model Architecture: A deep neural network is used as a model for architecture and has several layers, including an input layer, four hidden layers, and an output layer. Figure 2 shows the DNN model's architectural layout.

The Dense function, 96 neurons, kernel regularizer, and ReLU activation function define the input layer. The amount of features present in each input sample equals 46. Hence, 46 is the needed input shape. The model contains four hidden layers, each turned on by the ReLU function. The first hidden layer comprises 64 neurons, the second is the dropout layer, the third is 32 neurons, and the fourth layer is dropout with a value of 0.2. These layers help the model create complex representations from unstructured data. The output layer's single neuron represents each of the three classes in the classification task. The activation function used to determine these probabilities is sigmoid, and it is used to represent the model's prediction for each class for the input sample as a probability. By using the compile technique, the model is built. Because it excels in single-class classification scenarios, binary categorical cross-entropy is the loss function. The chosen optimizer, "Adam," is a well-known optimization tool. The accuracy of the model is assessed during the training phase.

C. EVALUATION MEASUREMENTS

Performance evaluation indicators are used in this study to evaluate a model's efficacy. Accuracy, precision, recall, the F1-score, and the confusion matrix are often used as evaluation measurements. A straightforward metric called accuracy counts how many instances out of all the examples are properly classified. To determine this value, apply equation 2 as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{2}$$

Precision is the percentage of times an estimate is right out of all the right times. To determine this value, apply equation 3 as follows:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

Recall measures the number of accurate positive predictions made, much like sensitivity. To determine this value, apply equation 4 as follows:

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

A well-rounded statistic, the F1-score takes the harmonic mean of precision and recall. To determine this value, apply equation 5 as follows:

$$F1 - score = 2 \times \frac{Precision + Recall}{Precision + Recall} \tag{5}$$

IV. EXPERIMENTAL ANALYSIS AND RESULT

The findings and results of examines using a federated learning and deep neural network model to classify attacks on IoT devices are provided here. 30% of the dataset is used for model testing, 10% for validation, and 60% for training. Hidden, DNN, and thick layers made up the architecture of the deep learning training system. We improved the model using the Adam optimizer and the binary cross-entropy loss function. The effectiveness of the model's predictions is evaluated using several metrics, including accuracy, precision, recall, F1-score, and the confusion matrix. Here is a detailed summary of the experimental results. The study's setup utilizes the Python 3.8 programming language and the Pycharm development framework. Writing, analyzing, and testing programs became simpler thanks to Pycharm, the integrated development environment (IDE) for Python. The most used operating system, Windows, supported Python programs with no issues. The popular machine learning programming language Python 3.8 provided several libraries and tools for data processing, analysis, and visualization.

A. EXPERIMENTAL RESULT OF THE SERVER SIDE

The Federated Learning (FL) system analysis focuses on the main parameter server and two clients. The server keeps track of every client while the Deep Neural Network model is being trained and gets model changes. The DNN model training process is recorded in the server-side data logs. Each of our trials is run three times to assure dependability. For the DNN study, we concentrate on the first three iterations of the experiments. This time frame is divided into two steps for each cycle: fitting and evaluation. During the fitting phase, the client transmits training data to the server; during the evaluating phase, both clients send test data to the server. Then, we compile all the data we have collected. Following the advised steps, we completed the experiment in 8.23 minutes.

B. EXPERIMENTAL RESULT OF CLIENT 1

The findings of a classification attack on IoT devices using a federated learning deep neural network on the CIC_IoT dataset are shown in Table 1.

TABLE 1. Proposed model result of client 1.

Rounds	Testing Accuracy	Precision	Recall	F1-score	Training Accuracy	Validation Accuracy
Round 1	99.00%	99.00%	99.00%	99.00%	98.79%	98.93%
Round 2	99.00%	99.00%	99.00%	99.00%	98.93%	99.05%
Round 3	99.00%	99.00%	99.00%	99.00%	99.04%	99.00%

During Round 1 of testing, the model correctly identified 99.00% of the attacks in the test set. The F1 score, recall, and precision were all 99.00%. While recall assesses how many positive attacks were favorably predicted, precision examines how many positive attacks were positive. The harmonic mean of recall and precision is used to calculate the F1-score, a comprehensive performance metric. The model did well on the training data, as seen by the first round's 98.79% training accuracy. The accuracy of the validation was 98.93%. In Round 2, the model demonstrated comparable performance, with values of 99.00% for testing accuracy, precision, and recall. The training accuracy increased to 98.93%, indicating that the model could learn more from the data. Additionally, the validation accuracy increased to 99.05%, showing improved generalization. Compared to Round 2, performance in Round 3 was essentially the same across all categories. The test's precision, recall, accuracy, and F1-score were 99.00%. Round 3's training accuracy is higher at 99.04% than round 2. However, round 3's validation accuracy dropped to 99.00% compared to round 2.

It can be seen from comparing the outcomes of each round that the model initially showed good accuracy and performance in Round 2. However, there was some variety in performance in the following rounds, with tinier changes in F1 score, accuracy, precision, and recall. Although there was some volatility in the validation accuracy, which suggests varied generalization ability, the model consistently maintained a high training accuracy, showing strong learning. The results demonstrate the great effectiveness of the federated learning approach for categorizing assaults on the CIC_IoT dataset. The system effectively categorizes IoT device threats by performing well compared to similar initiatives.

The visualization of the result is displayed in Figure 3, Figure 4, and Figure 5. For each round of the client 1 experiment, the training and validation accuracy, loss graphs, and confusion matrix were displayed. For the client 1 experiment, round 3 produced the best results. Figure 5a, Figure 5b illustrates a graph plotting the accuracy and loss for training and validation. Similarly, the validation loss is 0.057%, and the validation accuracy is 98.96%. The training loss is 0.065%, and the training accuracy is 98.93%. Figure 5c depicts the client 1 round 3 confusion matrix. The attack analysis outcomes from using a federated deep neural network on the CIC-IoT dataset, specifically for Client 1 during Round 3, correlate to the confusion matrix. 98.16% of the occurrences in the first row were correctly identified as positive attacks (true positives). On the other hand, the 1.84% of instances that were mistakenly projected as false positives revealed some misclassifications. Moving to the second row, we can observe that a very small percentage of incidents (0.05%) were falsely classified as negative attacks. 99.95% of the time, the incidents were accurately categorized as negative attacks (true negatives).

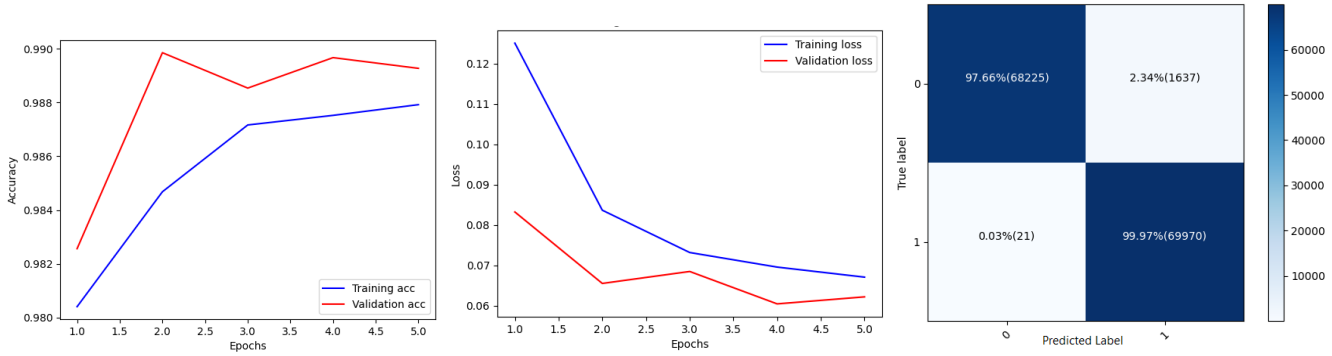
C. EXPERIMENTAL RESULT OF CLIENT 2

The findings of a classification attack on IoT devices using a federated learning deep neural network on the CIC_IoT dataset are shown in Table 2. During Round 1 of testing, the model correctly identified 99.00% of the attacks in the test set. The F1 score, recall, and precision were all 99.00%. While recall assesses how many positive attacks were favorably predicted, precision examines how many positive attacks were positive. The harmonic mean of recall and precision is used to calculate the F1-score, a comprehensive performance metric. The model did well on the training data, as seen by the first round's 98.79% training accuracy. The accuracy of the validation was 98.93%.

In Round 2, the model demonstrated comparable performance, with values of 99.00% for testing accuracy, precision, and recall. The training accuracy increased to 98.93%, indicating that the model could learn more from the data. Additionally, the validation accuracy increased to 99.05%, showing improved generalization. Compared to Round 2, performance in Round 3 was essentially the same across all categories. The test's precision, recall, accuracy, and F1-score were 99.00%. Round 3's training accuracy is higher at 99.04% than round 2. However, round 3's validation accuracy dropped to 99.00% compared to round 2.

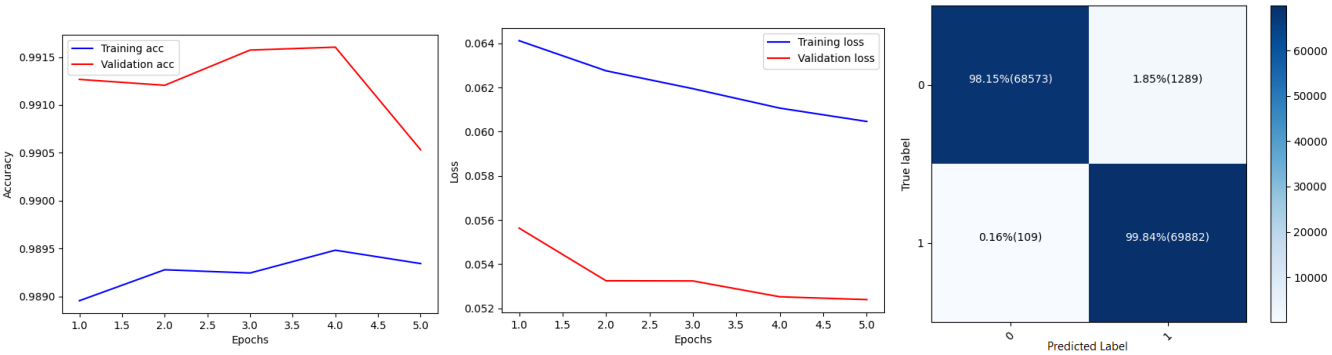
It can be seen from comparing the outcomes of each round that the model initially showed good accuracy and performance in Round 2. However, there was some variety in performance in the following rounds, with tinier changes in F1 score, accuracy, precision, and recall. Although there was some volatility in the validation accuracy, which suggests varied generalization ability, the model consistently maintained a high training accuracy, showing strong learning. The results demonstrate the great effectiveness of the federated learning approach for categorizing assaults on the CIC_IoT dataset. The system effectively categorizes IoT device threats by performing well compared to similar initiatives.

The visualization of the result is displayed in Figure 6, Figure 7, and Figure 8. For each round of the client 1 experiment, the training and validation accuracy, loss graphs, and confusion matrix were displayed. For the client 1



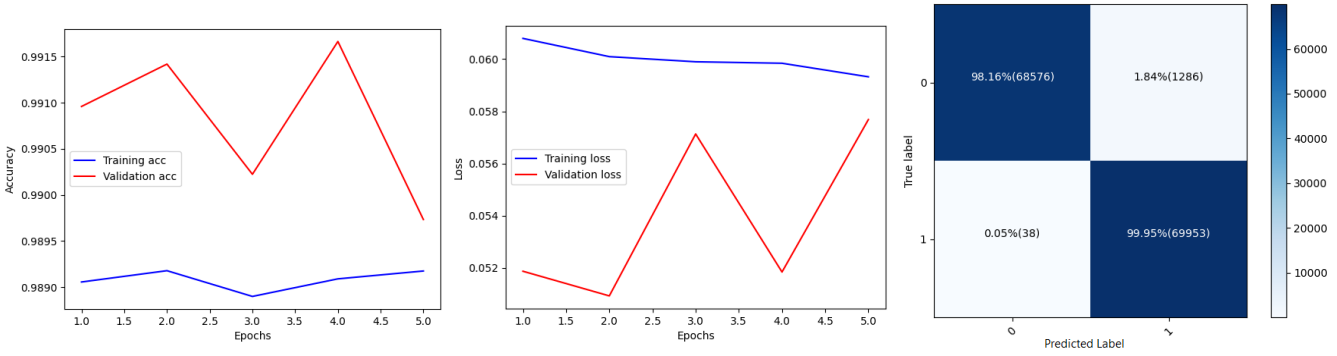
(a) Round 1 accuracy of validation and training (b) Round 1 loss of validation and training (c) Round 1 Confusion Metrics

FIGURE 3. Representation of client 1 round 1.



(a) Round 2 accuracy of validation and training (b) Round 2 loss of validation and training (c) Round 2 Confusion Metrics

FIGURE 4. Representation of client 1 round 2.



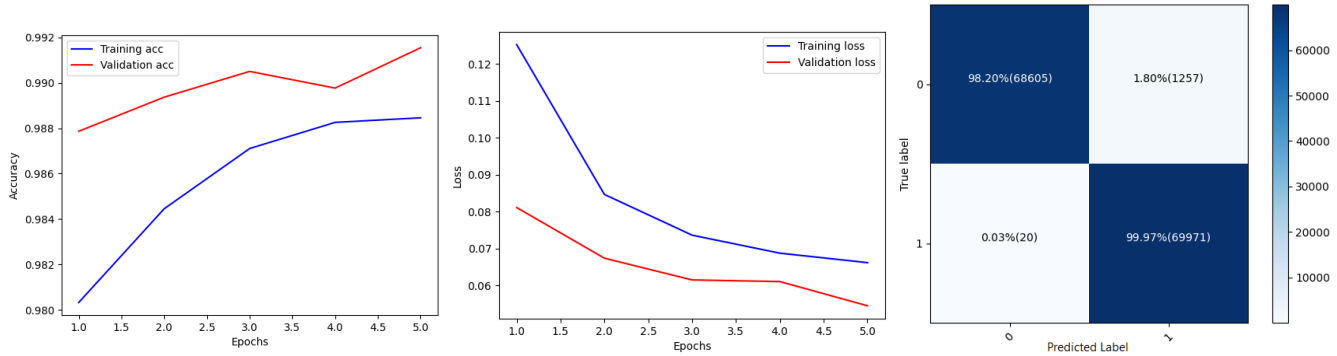
(a) Round 3 accuracy of validation and training (b) Round 3 loss of validation and training (c) Round 3 Confusion Metrics

FIGURE 5. Representation of client 1 round 3.

TABLE 2. Proposed model result of client 2.

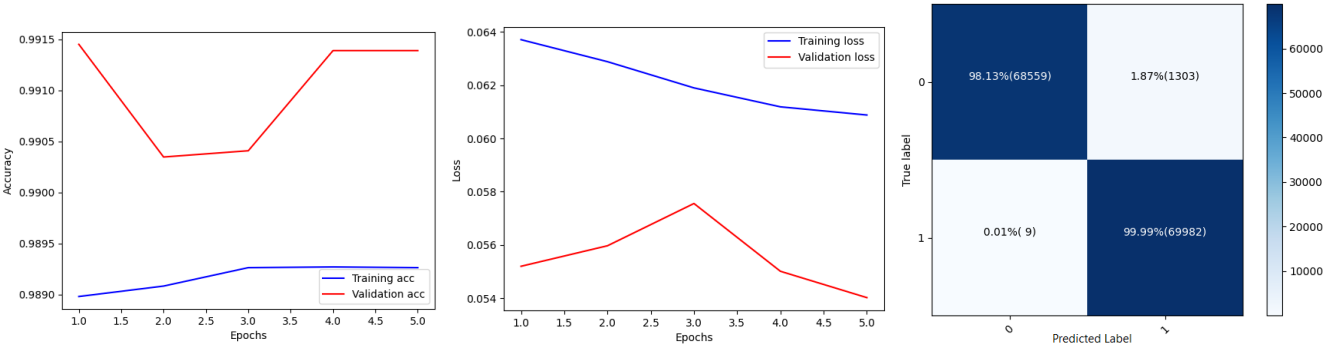
Rounds	Testing Accuracy	Precision	Recall	F1-score	Training Accuracy	Validation Accuracy
Round 1	99.00%	99.00%	99.00%	99.00%	98.85%	99.15%
Round 2	99.00%	99.00%	99.00%	99.00%	98.93%	99.14%
Round 3	99.00%	99.00%	99.00%	99.00%	99.04%	99.00%

experiment, round 2 produced the best results. Figure 7a and Figure 7b illustrates a graph plotting the accuracy and loss for training and validation. Similarly, the validation loss is 0.054%, and the validation accuracy is 99.15%.



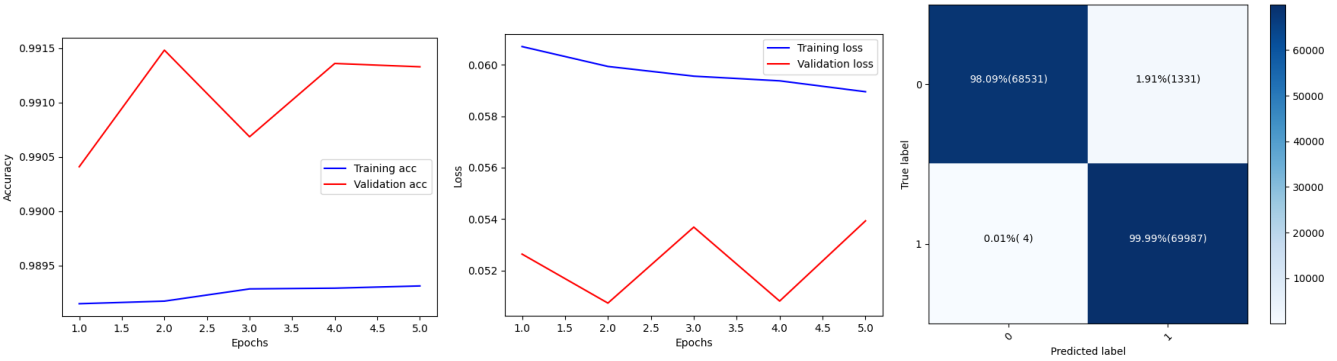
(a) Round 1 accuracy of validation and training (b) Round 1 loss of validation and training (c) Round 1 Confusion Metrics

FIGURE 6. Representation of client 2 round 1.



(a) Round 2 accuracy of validation and training (b) Round 2 loss of validation and training (c) Round 2 Confusion Metrics

FIGURE 7. Representation of client 2 round 2.



(a) Round 3 accuracy of validation and training (b) Round 3 loss of validation and training (c) Round 3 Confusion Metrics

FIGURE 8. Representation of client 2 round 3.

The training loss is 0.061%, and the training accuracy is 98.93%. Figure 7c depicts the client 1 round 2 confusion matrix. The attack analysis outcomes from using a federated deep neural network on the CIC-IoT dataset, specifically for client 1 during round 2, correlate to the confusion matrix. 98.13% of the occurrences in the first row were correctly identified as positive attacks (true positives). On the other hand, the 1.87% of instances that were mistakenly projected as false positives revealed some misclassifications. Moving to the second row, we can observe that a very small percentage of incidents (0.01%) were falsely classified as negative attacks. 99.99% of the time, the incidents were accurately categorized as negative attacks (true negatives).

V. CONCLUSION AND FUTURE SCOPE

This study presented a federated learning approach for identifying IoT device attacks. We used data preparation techniques (data balancing and feature normalization) to

prepare the data for model training to provide a trustworthy dataset for classification. A federated deep neural network model was trained using these features, making it possible to predict attacks accurately. The research produced a thorough and clean dataset that facilitated in-depth analysis and attack prediction. The suggested federated learning-based deep neural network attained an accuracy of 99.0%. In addition, we use federated learning techniques to improve the privacy and security of the training data. Additionally, to test the model's generalizability, our suggested approach can be expanded by employing additional datasets and adding attack features to the data. In the future, we plan on making new datasets with more IoT devices to prepare for future cybersecurity.

REFERENCES

[1] Z. A. El Houda, B. Brik, and S.-M. Senouci, "A novel IoT-based explainable deep learning framework for intrusion detection systems," *IEEE Internet Things Mag.*, vol. 5, no. 2, pp. 20–23, Jun. 2022.

[2] R. A. Devi and A. R. Arunachalam, "Enhancement of IoT device security using an improved elliptic curve cryptography algorithm and malware detection utilizing deep LSTM," *High-Confidence Comput.*, vol. 3, no. 2, Jun. 2023, Art. no. 100117.

[3] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1646–1685, 3rd Quart., 2020.

[4] Q. A. Al-Haija, M. Krichen, and W. A. Elhaija, "Machine-learning-based darknet traffic detection system for IoT applications," *Electronics*, vol. 11, no. 4, p. 556, Feb. 2022.

[5] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, and P. Faruki, "Network intrusion detection for IoT security based on learning techniques," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2671–2701, 3rd Quart., 2019.

[6] J. Zhang, "Distributed network security framework of energy internet based on Internet of Things," *Sustain. Energy Technol. Assessments*, vol. 44, Apr. 2021, Art. no. 101051.

[7] J. Jeon, B. Jeong, S. Baek, and Y.-S. Jeong, "Hybrid malware detection based on bi-LSTM and SPP-Net for smart IoT," *IEEE Trans. Ind. Informat.*, vol. 18, no. 7, pp. 4830–4837, Jul. 2022.

[8] F. Alasmary, S. Alraddadi, S. Al-Ahmadi, and J. Al-Muhtadi, "Shield-RNN: A distributed flow-based DDoS detection solution for IoT using sequence majority voting," *IEEE Access*, vol. 10, pp. 88263–88275, 2022.

[9] A. Pinhero, M. Anupama, P. Vinod, C. A. Visaggio, N. Aneesh, S. Abhijith, and S. AnanthaKrishnan, "Malware detection employed by visualization and deep neural network," *Comput. Secur.*, vol. 105, Jun. 2021, Art. no. 102247.

[10] L. Yi, X. Deng, L. T. Yang, H. Wu, M. Wang, and Y. Situ, "Reinforcement-learning-enabled partial confident information coverage for IoT-based bridge structural health monitoring," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3108–3119, Mar. 2021.

[11] S. K. Smmarwar, G. P. Gupta, and S. Kumar, "A hybrid feature selection approach-based Android malware detection framework using machine learning techniques," in *Cyber Security, Privacy and Networking*. Cham, Switzerland: Springer, 2022, pp. 347–356.

[12] A. Rahman, K. Hasan, D. Kundu, M. J. Islam, T. Debnath, S. S. Band, and N. Kumar, "On the ICN-IoT with federated learning integration of communication: Concepts, security-privacy issues, applications, and future perspectives," *Future Gener. Comput. Syst.*, vol. 138, pp. 61–88, Jan. 2023.

[13] P. Ruzafa-Alcázar, P. Fernández-Saura, E. Mármol-Campos, A. González-Vidal, J. L. Hernández-Ramos, J. Bernal-Bernabe, and A. F. Skarmeta, "Intrusion detection based on privacy-preserving federated learning for the industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 1145–1154, Feb. 2023.

[14] K. Kostas, M. Just, and M. A. Lones, "Externally validating the IoTDevID device identification methodology using the CIC IoT 2022 dataset," 2023, *arXiv:2307.08679*.

[15] F. S. D. L. Filho, F. A. F. Silveira, A. de Medeiros Brito Junior, G. Vargas-Solar, and L. F. Silveira, "Smart detection: An online approach for DoS/DDoS attack detection using machine learning," *Secur. Commun. Netw.*, vol. 2019, pp. 1–15, Oct. 2019.

[16] J. J. Q. Yu, Y. Hou, and V. O. K. Li, "Online false data injection attack detection with wavelet transform and deep neural networks," *IEEE Trans. Ind. Informat.*, vol. 14, no. 7, pp. 3271–3280, Jul. 2018.

[17] M. Sarhan, S. Layeghy, and M. Portmann, "Feature analysis for machine learning-based IoT intrusion detection," 2021, *arXiv:2108.12732*.

[18] O. Friha, M. A. Ferrag, L. Shu, L. Maglaras, K.-K.-R. Choo, and M. Nafaa, "FELIDS: Federated learning-based intrusion detection system for agricultural Internet of Things," *J. Parallel Distrib. Comput.*, vol. 165, pp. 17–31, Jul. 2022.

[19] F. Naeem, M. Ali, and G. Kaddoum, "Federated-learning-empowered semi-supervised active learning framework for intrusion detection in ZSM," *IEEE Commun. Mag.*, vol. 61, no. 2, pp. 88–94, Feb. 2023.

[20] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, and A. A. Ghorbani, "CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment," *Sensors*, vol. 23, no. 13, p. 5941, Jun. 2023.

[21] A. Fernandez, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary," *J. Artif. Intell. Res.*, vol. 61, pp. 863–905, Apr. 2018.

[22] H. Henderi, "Comparison of min-max normalization and Z-score normalization in the K-nearest neighbor (kNN) algorithm to test the accuracy of types of breast cancer," *IJIIS, Int. J. Informat. Inf. Syst.*, vol. 4, no. 1, pp. 13–20, Mar. 2021.

SIDRA ABBAS received the B.S. degree from the Department of Computer Science, COMSATS University Islamabad, Islamabad, Pakistan. Her research interests include but are not limited to computer forensics, machine learning, criminal profiling, software watermarking, intelligent systems, and data privacy protection.

ABDULLAH AL HEJAILI received the bachelor's degree in computer science from the Tabuk Teachers College, Saudi Arabia, in 2007, and the master's degree in computer science from CLU, USA, in 2011. He is currently pursuing the Ph.D. degree with the Informatics School, University of Sussex. He is also a Lecturer of computer science with the University of Tabuk. His research interests include technology-enhanced learning, image processing, virtual, augmented reality, motion capture, and education applications.

GABRIEL AVELINO SAMPEDRO (Member, IEEE) received the B.S. and M.S. degrees in computer engineering from Mapúa University, Manila, Philippines, in 2018, and the Ph.D. degree in IT convergence engineering from the Kumoh National Institute of Technology, in 2023. He is currently an Assistant Professor with the Faculty of Information and Communication Studies, University of the Philippines Open University, and a Researcher with the Center for Computational Imaging and Visual Innovations, De La Salle University. His research interests include real-time systems, embedded systems, robotics, and biomedical engineering.

MIDETH ABISADO is currently with the College of Computing and Information Technologies, National University, Philippines. Her research interests include deep learning (artificial intelligence), learning (artificial intelligence), the IoT, convolutional neural nets, feature extraction, mobile computing, recurrent neural nets, social networking (online), three-dimensional printing, and 5G mobile.

AHMAD S. ALMADHOR (Member, IEEE) received the B.S.E. degree in computer science from Jouf University (formerly Al-Jouf College), Al-Jouf, Saudi Arabia, in 2005, the M.E. degree in computer science and engineering from the University of South Carolina, Columbia, SC, USA, in 2010, and the Ph.D. degree in electrical and computer engineering from the University of Denver, Denver, CO, USA, in 2019. From 2006 to 2008, he was a Teaching Assistant and a College of Sciences Manager, and then a Lecturer with Jouf University, from 2011 to 2012. Then, he became a Senior Graduate Assistant and a Tutor Advisor with the University of Denver, in 2013 and 2019. He is currently an Assistant Professor of CEN and the VD of the Computer and Information Science College, Jouf University. His research interests include AI, blockchain, networks, smart and microgrid cyber security, integration, image processing, video surveillance systems, PV, EV, machine, and deep learning. His awards and honors include the Jouf University Scholarship (Royal Embassy of Saudi Arabia in D.C.) and the Al-Jouf Governor Award for excellence.

TARIQ SHAHZAD received the B.E. and M.S. degrees from COMSATS University Islamabad, Sahiwal, Pakistan, in 2006 and 2014, respectively, and the Ph.D. degree from the University of Johannesburg, South Africa, in 2021. He is currently a Research Fellow with the University of Johannesburg. He has published several research articles in journals and international conferences. His research interests include biomedical signals processing, machine learning, and computer vision.

KHMAIES OUAHADA was the Head of the Department, from December 2018 to December 2021. He is currently a Full Professor with the Department of Electrical and Electronic Engineering Science, Faculty of Engineering and the Built Environment. He lectures signal processing 4A after teaching two modules, telecommunications 3B, and signal processing 3B, which form a critical foundation for fourth-year Exit Level Modules. In the past, he has taught other introductory and advanced modules and has done an excellent job organizing and lecturing. His lecture notes for electrotechnics 2A, electrotechnics 2B, and signals and systems 3A are still used by his successors. He was awarded the Vice-Chancellors Teaching and Learning Excellence Award, in 2016. This is a natural result of his hard work and passion as a lecturer.