IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 11 May 2023, accepted 24 May 2023, date of publication 29 May 2023, date of current version 8 June 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3280798

RESEARCH ARTICLE

Blockchain Assisted Data Edge Verification With Consensus Algorithm for Machine Learning Assisted IoT

THAVAVEL VAIYAPURI[1], (Member, IEEE), K. SHANKAR[2,3], (Senior Member, IEEE), SURENDRAN RAJENDRAN[2], (Member, IEEE), SACHIN KUMAR[3], (Senior Member, IEEE), SRIJANA ACHARYA[4], AND HYUNIL KIM[4]

[1]College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
[2]Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai 602105, India
[3]Big Data and Machine Learning Laboratory, South Ural State University, 454080 Chelyabinsk, Russia
[4]Department of Convergence Science, Kongju National University, Chungcheongnam-do, Gongju-si 32588, South Korea

Corresponding author: Hyunil Kim (hyunil89@kongju.ac.kr)

This work was supported in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) Grant funded by the Korea Government (MSIT) (Development of User Identity Certification and Management Technology for Self-Sovereign Identity Applications and Robust AI and Distributed Attack Detection for Edge AI Security) under Grant 2021-0-00565 and Grant 2021-0-00511.

ABSTRACT Internet of Things (IoT) devices are becoming increasingly ubiquitous in daily life. They are utilized in various sectors like healthcare, manufacturing, and transportation. The main challenges related to IoT devices are the potential for faults to occur and their reliability. In classical IoT fault detection, the client device must upload raw information to the central server for the training model, which can reveal sensitive business information. Blockchain (BC) technology and a fault detection algorithm are applied to overcome these challenges. Generally, the fusion of BC technology and fault detection algorithms can give a secure and more reliable IoT ecosystem. Therefore, this study develops a new Blockchain Assisted Data Edge Verification with Consensus Algorithm for Machine Learning (BDEV-CAML) technique for IoT Fault Detection purposes. The presented BDEV-CAML technique integrates the benefits of blockchain, IoT, and ML models to enhance the IoT network's trustworthiness, efficacy, and security. In BC technology, IoT devices that possess a significant level of decentralized decision-making capability can attain a consensus on the efficiency of intrablock transactions. For fault detection in the IoT network, the deep directional gated recurrent unit (DBiGRU) model is used. Finally, the African vulture optimization algorithm (AVOA) technique is utilized for the optimal hyperparameter tuning of the DBiGRU model, which helps in improving the fault detection rate. A detailed set of experiments were carried out to highlight the enhanced performance of the BDEV-CAML algorithm. The comprehensive experimental results stated the improved performance of the BDEV-CAML technique over other existing models with maximum accuracy of 99.6%.

INDEX TERMS Blockchain, Internet of Things, consensus algorithm, fault detection, deep learning, hyperparameter tuning.

I. INTRODUCTION

With the advent of 5G, wireless sensor networks (WSNs), and relevant technology, the Internet of Things (IoT) has become more prominent as a new concept to satisfy the requirements of ubiquitous, flexible, and agile availability of cyberspace from physical systems [1]. But the present centralized IoT framework is greatly limited by security challenges, single point of failure, robustness, and data privacy [2]. Nowadays, blockchain (BC) is a potential solution to sort out such issues

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

because of its capability to maintain an immutable open ledger that is easily accessed by all but is tamper-proof [3]. As well a wide range of innovative IoT applications is enabled by the rapid development of edge computing. Few research works are done on this subject, and various tools, like BC and IoT, are utilized to make an intelligent supply chain. Still, research gaps in this area are found, like the study of relations of these three areas: BC, smart supply chain, and IoT [4]. Another application is the absence of categorization of IoT features and BC that disturbs building an intelligent supply chain. BC is a unique e-book, a kind of data and reporting mechanism that records its worth [5]. The only difference between BC and other mechanisms is that the saved data is shared between network members. Encryption usage makes it almost impossible to manipulate or delete the recorded data.

BC technology presents immense benefits like stability, trust, speed, effectiveness, precision and independence [6]. The unique features of a BC program (unchangeable records, peer-to-peer relationship, approval schedule, and autonomous) lead to an error elimination of transfer, increased productivity, the use of data, data security assurance and time and money saving [7]. Decentralization indicates the absence of central authority or intermediaries; the selected or individual participants in one BC can validate the reports of their trading partners and be suitable to use the whole database and entire history without the intermediary's help [8]. Further, by eliminating the role of management intermediary, BC eradicates the necessities of centralized management; the distributed system called Peer-to-peer systems composed of personal computers, which offer their computational resources (information distribution, storage capacity) directly to others [9]. Indeed, BC technology is devised in dispersed systems to alleviate any single point of failure; customers need not believe their service providers. In the industrial Internet of Things (IIoT), one significant problem is Fault detection. In traditional device failure recognition of IIoT, for centralised model training, client devices should upload local raw information to a central server [10]. This may cause problems like data privacy since the local data of clients can be business sensitive. For instance, using air conditioning in hotels possibly reflects the occupancy rate.

This study develops a new Blockchain Assisted Data Edge Verification with Consensus Algorithm for Machine Learning (BDEV-CAML) technique for IoT Fault Detection purposes. In BC technology, IoT devices that possess a significant level of decentralized decision-making ability can reach an agreement regarding the effectiveness of intrablock transactions. The deep directional gated recurrent unit (DBiGRU) model is used for fault detection in the IoT network. Finally, the African vulture optimization algorithm (AVOA) technique is utilized for the optimal hyperparameter tuning of the DBiGRU model, which helps in improving the fault detection rate. A detailed set of experiments were carried out to highlight the enhanced performance of the BDEV-CAML algorithm.

II. RELATED WORKS

Trivedi et al. [11] devised a DL and BC-based EV fault detection (FD) structure to find different kinds of faults: battery faults, air tire pressure, and vehicle temperature. Further, to achieve the FD data transaction with high reliability and scalability for EV, the author uses a 5G wireless network, including an interplanetary file system (IPFS). Initially, the author uses an LSTM and a CNN technique to deal with battery fault detection, air tire pressure fault, and anomaly detection for temperature fault to forecast the faulty dataset, guaranteeing users a safer journey. In [12], the author target to detect a technique for potential FD in IoT gadgets. Based on BC, an IoT network architecture is initially developed, and a data edge authentication system is designed; the BC is utilized for assuring that data could not be tampered with and the precision can be guaranteed. Eventfully, a dataset accuracy-weighted RF-related PSO was devised.

Belhadi et al. [13] developed an innovative structure to precisely find anomalous patterns in privacy RL in a heterogeneous and distributed energy environment. The local outlier factor was accomplished for deriving local abnormal patterns in all sites of the distributed energy environment. With BC technology, reinforcement privacy learning combines local anomalous patterns into global complicated ones. Huang et al. [14] presented a Gaussian Bernoulli restricted Boltzmann machine (GBRBM)-oriented DNN innovative algorithm for transforming the FD into a classifier issue. The presented technique outpaces other baseline ML techniques by real trace-driven experiments. In [15], an innovative Energy-Efficient Heterogeneous Fault Management system was devised to achieve such heterogeneous faults in IWSN. The three new diagnosis techniques can achieve efficient heterogeneous FD in this technique. The Tuned SVM classifier facilitated the classification of the heterogeneous faults where the tuning parameter of the presented method is optimized using the Hierarchy-related Grasshopper Optimization method.

Mittra et al. [16] emphases on investigating these main issues in the secondary transmission system, proposing techniques to implement and integrate modern technologies like BC and IoT, and sightseeing the scope of enhancements in the current system. The research workers have modernised conventional techniques and re-applied devices and equipment. Zhang et al. [17] present a new adaptive privacy-preserving federated learning method called AdaPFL for FD in IoS, which organizes various shipping agents to develop a technique by sharing model parameters without data leakage. Initially, the author uses two tasks as instances to show that a smaller portion of model parameters may expose the raw data of shipping agents. To protect the basic information of shipping agents, the Paillier-based communication technique was devised.

In [18], a maintenance-oriented KG is presented initialy dependent upon a explicit domain oriented ontology scheme and accumulated preserved data. Afterward, an Attention-Based Compressed Relational GCN was presented for predicting possible solutions and describe fault in preserve tasks. Xia et al. [19] presents a novel hypergraph convolution network (HGCN) based model to forecast MRR from the CMP procedure. A major contributions contains: (1) a generic hypergraph method for representing the interactions of difficult tools; and (2) a temporal enabled forecast method for learning the difficult data correlation and high order representation dependent upon the hypergraph. In [20], a vertical federated learning (FL) method, privacy-preserving boosting tree was established to collaborative fault analysis of industrial practitioners but maintained anonymity.

Xia et al. [21] examines a residual-HGCN (Res-HGCN) technique which holistically drive in tools structure and operational processes as a hypergraph procedure as data-driven method allowing for reaction amongst equipment's mechanisms. Keung et al. [22] introduces the execution of developing ARP for IIoT and resource synchronisation flexible robotic and facility control scheme for addressing this challenge. In [23], the authors propose for addressing the value construction of exploiting the IIoT-driven resource synchronization and sharing-based robotic mobile fulfillment system (RMFS) for enhancing the entire operational efficiency and effectiveness in the data transmission and synchronization of resources.

The number of parameters of DL methods increases quickly because of the continuous deepening of the method that give rise to model overfitting. Meanwhile, various hyperparameters have a huge impact on the ability of the CNN method. To be specific, the hyperparameters like learning rate selection, epoch count, and batch size are vital to get productive results. As the trial and error method for hyperparameter tuning is a erroneous and tedious process, metaheuristic methods are adopted. Hence, the authors have used AVOA method for the parameter selection of the DBiGRU approach in this study.

III. THE PROPOSED MODEL

In this study, we have designed a novel BDEV-CAML algorithm to identify faults in the IoT environment effectively. The BDEV-CAML technique integrated the advantages of the blockchain, IoT, and ML concepts for boosting the IoT network's trustworthiness, efficacy, and security. In the context of intrablock transactions, IoT devices with strong decentralized decision-making abilities can agree on the effectiveness. Fig. 1 represents the overall procedure of the BDEV-CAML approach.

A. BLOCKCHAIN TECHNOLOGY

Blockchain (BC) technology is the eminent domain for safety and trust, which can be applied to any relevant topic to keep data and information private [24]. Likewise, it is a groundbreaking technology for distributed and decentralized computing frameworks that support the data with encrypted blocks in the chain. Digital data related to time, date amount, transactions, and so on are enlarged in the transaction process and stored in the block. The stored information is now available within the distributed network, with a participant node to authenticate the transaction. Every node in BC is connected and assists the transaction and crypto code. An additional feature in BC technology is the mathematical algorithm that is extremely powerful in these networks. It is responsible for blocking authentication to minor nodes without affecting the data; for that reason, BC is transparent and secure. There exist recommendation systems based on BC and knowledge discovery technology and many research requirements for tackling security challenges. This procedure must implement the incorporation of IoT and BC. Likewise, the security challenges, which the research workers state, make the BC a better choice. BC's main characteristics are programmability, trust, security, and so on. A BC comprises a private BC, a consortium BC, or a public BC. The public BC is well-known for digital currency. The primary goal of the consortium BC is to integrate the service trading and stakeholder entity.

B. IoT NETWORK MODEL AND DISTRIBUTED LEDGER DEPLOYMENT

If the edge gateway fails, IoT using single-edge gateways is prone to single-point failure since the whole IoT will be disconnected. Multiple-edge gateways decrease the delay and distance of communication of information and prevent energy loss based on the fast consumption of the IoT node nearby the edge gateway. In the presented method, multiple edge gateways are the edge server and are applied to authenticate the information received from the BC [12]. The edge node process the received statistics and encapsulate the processed information, timestamp, and other transaction data into the block that can be represented as the data block. Data authentication can be accomplished by transporting these blocks. Each edge gateway retains a synchronized and shared distributed ledge. Every transaction between the edge gateways is stored in the ledger, like data exchange or asset records. Due to the computing power, limited storage capacity, energy, etc., the distributed ledger is positioned on edge gateways and is retained by the edge gateway. The information transferred in IoT is stored in the ledger with decentralised features, guaranteeing that the information is tamper-proof.

Data Consensus Algorithm: In BC technology, nodes equipped with extensive decentralized decision-making abilities could attain a consensus on the efficiency of intrablock transactions. To guarantee consensus amongst edge gateways, the conventional BC technique depends mainly on the computing power of distributed edge gateway. With comparatively poor computation power, this doesn't apply to IoT. Consequently, the study developed a data consensus model of IoT. The hash function is performed for transforming the information into respective hash values once the edge gateway receives the information [12]. Every evaluated hash value corresponds to the data point, and the new information could not be recovered through hashing. The edge gateway receives the information store, the destination edge gateway, and other related data in the block. The data block is transmitted to each edge gateway for authentication and waits for the authentication outcomes. The information is authenticated by checking the ledger once the edge gateway receives the information from other edge gateways. The edge gateway waits for confirmation and then gives feedback on the authentication outcome (correct or incorrect). The received information is processed based on the authentication result from diverse edge gateways accountable for confirmation. Based on BC technology, if the majority of the confirmed edge gateway passes authentication (the confirmation outcome is true), the information is labelled as correct otherwise, they are labelled as incorrect. Lastly, the processed information is transmitted to the management system to detect faults.

C. FAULT DETECTION USING DBIGRU MODEL

For fault detection in the IoT network, the AVOA with DBiGRU technique is utilized. Unlike the standard unidirectional RNNs, Bi-directional network includes two single hidden RNN layers (the backward and forward layers) in its framework [25]. Every layer is interconnected with input and output layers, correspondingly. The bidirectional model enables its network to learn the tourist volume sequence from the future and past directions. Backward and forward layers in the network read input series $x(x_1, x_2, x_3, \ldots, x_{n-1}, x_n)$ from two opposite directions, in which $x^{forward} = (x_1, x_2, x_3, \ldots, x_{n-1}, x_n)$ and $x^{backward} = (x_n, x_{n-1}, \ldots x_t \ldots, x_2, x_1)$, then attain a forward hidden state $\vec{h}_t(\vec{h}_1, \vec{h}_2, \ldots, \vec{h}_{n-1}, \vec{h}_n)$ and backward hidden state $\overleftarrow{h}_t(\overleftarrow{h}_1, \overleftarrow{h}_2, \ldots, \overleftarrow{h}_{n-1}, \overleftarrow{h}_n)$, correspondingly (Eqs. (1) and (2)). Consequently, backwards and forward series are concentrated and produce the final output series $y(y_1, y_2, \ldots y_t \ldots, y_{n-1}, y_n)$ that can be evaluated by the following expression.

$$\vec{h}_t = f\left(w_{x\vec{h}} \cdot x_t + W_{\vec{h}\vec{h}} \cdot \vec{h} + b_{\vec{h}}\right) \tag{1}$$

$$\overleftarrow{h}_t = f\left(W_{x\overleftarrow{h}} \leftarrow \cdot x_t + W_{\overleftarrow{h}\overleftarrow{h}} \cdot h_{t+1} + b_{\overleftarrow{h}}\right) \tag{2}$$

$$y_t = w_{y\vec{h}} \cdot \vec{h}_t + w_{y\overleftarrow{h}} \cdot \overleftarrow{h}_t + b_y \tag{3}$$

where $b_{\overleftarrow{h}}$, and b_y denotes the respective bias vector, $\vec{h}_t$, $\overleftarrow{h}_t$ denote forward and backward propagation, correspondingly; f represents a nonlinear activation function (viz., sigmoid function); and $w_{x\vec{h}}$, $W_{\vec{h}\vec{h}}$, $W_{x\overleftarrow{h}}$, $W_{\overleftarrow{h}\overleftarrow{h}}$, $W_{\overleftarrow{h}\overleftarrow{y}}$, and $W_{x\overleftarrow{h}}$ characterize the respective weight coefficient.

GRU cell was employed for adding to the abovementioned bi-directional network that is a variant of LSTM cell and called an improved version of RNN cell. Though GRU was barely used for tourist volume prediction, it accomplished the desired forecast effects the same as LSTM in other time-series forecasts. GRU simplifies the gating model from the 3 LSTM gates: forget, input, and output. A standard GRU cell comprises reset and update gates to decrease the computational cost. Fig. 2 signifies the structure of GRU. The reset gate defines what data the existing step could be accessed from h_{t-1} and x_t, which can be evaluated by Eq. (4). Consequently, a candidate vector, $\tilde{h}_t$, is generated using tanh function where the output of reset gate, r_t, only influence h_{t-1}. Moreover, the update gate controls the influence of the preceding state, h_{t-1}, and candidate vector, $\overline{h}$, on state vector, h_t, evaluated by Eqs. (6) & (7).

$$r_t = \sigma\left(W_r \cdot \left[h_{t-1'} x_t\right] + b_r\right) \tag{4}$$

$$\tilde{h}_t = tanh\left(W_h \cdot \left[r_t O h_{t-1}, x_t\right] + b_h\right) \tag{5}$$

$$u_t = \sigma\left(W_u \cdot \left[h_{t-1'} x_t\right] + b_u\right) \tag{6}$$

$$h_t = (1 - u_t)\, h_{t-1} + u_t \tilde{h}_t \tag{7}$$

where h_{t-1} and h_t epitomize prior cell state and existing cell state; $\tilde{h}_t$ represents a candidate activation vector; r_t and u_t characterize the output of reset and update gates, correspondingly; and b_r, b_h, and b_u denote bias vectors; W_r, W_h, and W_u characterize weight matrixes, $\odot$ signifies the Hadamard product.

D. HYPERPARAMETER TUNING USING AVOA

Finally, the AVOA technique is utilized for the optimal hyperparameter tuning of the DBiGRU model, which helps improve the fault detection rate. The hyperparameters involved are learning rate, batch size, and number of epochs. The AVOA is motivated by the navigational and foraging behaviours of African vultures [26]. The biological nature of vultures with reference to competing and searching for food is considered in four distinct phases. Consider N vultures in the environment representing the population $n = \{1, 2, \ldots, N\}$. Next, calculate the fitness function of every location position. Where p_n is the probability of choosing the 1st or 2nd group that can be evaluated by

$$p_n = \frac{F_n}{\sum_{n=1}^{N} F_n} \tag{8}$$

In Eq. (8), F_n denotes the fitness function of the *nth* location. Next, the construction of the 1st and 2nd groups in every iteration can be attained using Eq. (9):

$$R\,(it) = \begin{cases} \textit{first group}, & p_n = L_1 \\ \textit{second group}, & p_n = L_2 \end{cases} \tag{9}$$

The satiated vulture with adequate energy can be moved to longer distances for searching food, while the hungry one cannot. Where the ranges of L_1 and L_2 are $0 \leq L_1, L_2 \leq 1$ and $L_1 + L_2 = 1$. The rate of being hungry or satiated defines the movement from the exploration stage to the exploitation stage that can be formulated by using Eq. (10) and (11):

$$A = (2 \times rand_1 + 1) \times x \times \left(1 - \frac{it}{IT_{\max}}\right) + y \tag{10}$$

$$y = h \times \left(\sin^z\left(\frac{\Pi}{2} \times \frac{it}{IT_{\max}}\right) + \cos\left(\frac{\Pi}{2} \times \frac{it}{IT_{\max}}\right) - 1\right) \tag{11}$$

where x, h, and $rand_1$ denote the random integer differs from -1 to 1, -2 to 2, and 0 to 1, correspondingly; A represents the vulture with the highest energy, z describes the prospect of entering the exploration phase; it and $IT_{\max}$ denote the current and maximal iteration, correspondingly. The exploration phase defines the process of finding food by the African vultures where the parameter $p_1, 0 \leq p_1 \leq 1$ determines the selection of strategy thus,

$$P\left(it+1\right)=\begin{cases}(6)\,, & p_1 \geq rand_2\\ (8)\,, & p_1 < rand_2\end{cases} \tag{12}$$

$$P\left(it+1\right)=R\left(it\right)-D\left(it\right)\times A \tag{13}$$

$$D\left(it\right)=\left|q\times R\left(it\right)-P\left(it\right)\right| \tag{14}$$

Here P(it + l) indicate the vultures' location vector in the following iteration. A and $R(it)$ are attained, correspondingly $q = 2 \times rand_3$, where $rand_3$ indicates a random integer ranging from zero to one.

$$P\left(it+l\right)=R\left(it\right)-A+rand_4\times\left(\left(u_b-\mathrm{l}\right)\times rand_5+\mathrm{l_b}\right) \tag{15}$$

Eq. (15), u_b and 1_b denote the upper and lower boundaries of the parameter; correspondingly, $Rand_4$ and $Rand_5$ show the random integers within [0, 1]. The exploitation phase has two distinct stages. The selection of any strategy relies on the variables p_2 and p_3. p_2 and p_3 values lie within [0, 1]. If $|F|$ lies within [1, 0.5], then the exploitation phase enters the initial stage that defines two strategies, namely siege fight and rotating flight:

$$P\left(it+1\right)=\begin{cases}(10)\,, & p_2 \geq rand_6\\ (11)\,, & p_2 < rand_6\end{cases} \tag{16}$$

In Eq. (16), $rand_6$ denotes a random integer within [0, 1]. The solution to Eq. (16) is shown below:

$$P\left(it+1\right)=D\left(it\right)\times\left(A+rand_7\right)-d\left(it\right) \tag{17}$$

$$D\left(it\right)=R\left(it\right)-P\left(it\right) \tag{18}$$

Next, the rotational flight of the vulture can be modelled by:

$$P\left(it+l\right)=R\left(it\right)-\left(M_1+M_2\right) \tag{19}$$

$$M_1=R\left(it\right)\times\left(\frac{rand_8\times P\left(it\right)}{2\pi}\right)\times\cos\left(P\left(it\right)\right) \tag{20}$$

$$M_2=R\left(it\right)\times\left(\frac{rand_9\times P\left(it\right)}{2\pi}\right)\times\sin\left(P\left(it\right)\right) \tag{21}$$

where $rand_8$ and $rand_9$ denote two random integers within [0, 1]. if $|F| > 0.5$, then the exploitation stage enters the next stage that describes the two dissimilar approaches of vultures, namely aggressive and accumulation fight and siege to search for food. Based on the following condition, the selection of any strategy can be done:

$$P\left(it+1\right)=\begin{cases}(16)\,, & p_3 \geq rand_7\\ (19)\,, & p_3 < rand_7\end{cases} \tag{22}$$

where

$$P\left(it+1\right)=\frac{B_1+B_2}{2} \tag{23}$$

TABLE 1. Precision rate analysis of BDEV-CAML approach with other systems under varying classes.

Precision Rate (%)				
Class	BDEV-CAML	PSO-DAWRF	DAWRF	RF
CPUHog	99.55	98.6	97.06	95.87
MemoryOF	99.24	98.1	96.98	95.52
Scanning	99.57	98.33	97.17	96.04
IOHog	99.72	98.74	97.56	95.99
DOS	99.41	97.91	96.41	95.25

$$B_1=Best_{vulture1}\left(it\right)-\frac{Best_{vulture1}\left(it\right)\times P\left(it\right)}{Best_{vulture1}(it)\times P(it)^2}\times A \tag{24}$$

$$B_2=Best_{vulture2}\left(it\right)-\frac{Best_{vulture2}\left(it\right)\times P\left(it\right)}{Best_{vulture2}(it)\times P(it)^2}\times A \tag{25}$$

where $Best\_vulture1(it)$ and $Best\_vulture2(it)$ represent the optimum vulture of the 1st and 2nd groups, correspondingly, in the existing iteration. The aggressive competition amongst the vultures is given below:

$$P(it+1)=R(it)-|d(it)|\times A\times Levy(d) \tag{26}$$

where d_1 signifies the dimension of optimization problem:

$$Levy(x)=0.01\times\frac{u\times\sigma}{|v|^{\frac{1}{\beta'}}},$$

$$\sigma=\left(\frac{Y\left(1+\beta\right)\times\sin\left(\frac{\pi\beta}{2}\right)}{Y\left(1+2\beta\right)\times\beta\times 2\left(\frac{\beta-1}{2}\right)}\right)^{\frac{1}{\beta}} \tag{27}$$

Fitness selection is a critical factor in the AVOA method. Solution encoding is used to assess the goodness (aptitude) of the solution candidate. Then, the accuracy value is the primary condition applied to design a fitness function.

$$Fitness=\max\left(P\right) \tag{28}$$

$$P=\frac{TP}{TP+FP} \tag{29}$$

From the expression, TP characterizes the true positive, and FP indicates the false positive value.

IV. RESULTS AND DISCUSSION

In this section, the experimental outcomes of the BDEV-CAML technique are studied under different measures. The results are inspected under various types of faults. In Table 1, and Fig. 3, a comparative precision rate (PR) analysis of the BDEV-CAML technique is given.

The results indicate that the RF model attains the least efficiency, whereas the PSO-DAWRF and DAWRF models attain closer results. Nevertheless, the BDEV-CAML technique reaches better PR values. For instance, with CPUHog fault class, the BDEV-CAML technique obtains increasing PR of 99.55% while the PSO-DAWRF, DAWRF, and RF

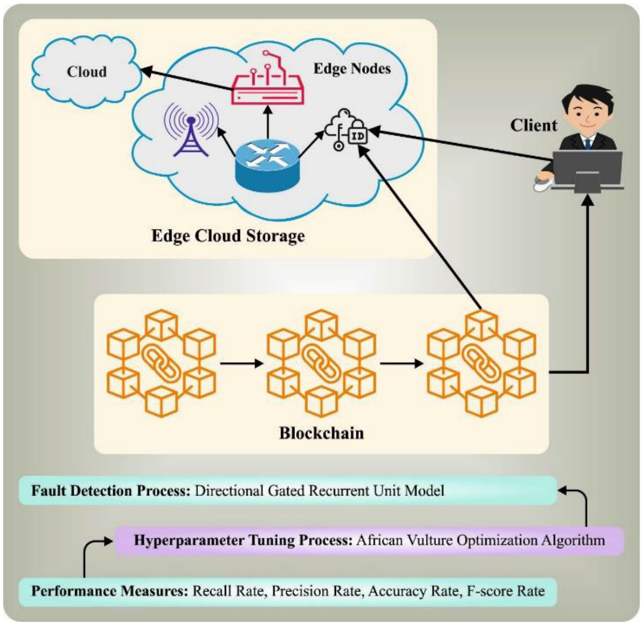

Cloud
Edge Nodes
Client
Edge Cloud Storage
Blockchain
Fault Detection Process: Directional Gated Recurrent Unit Model
Hyperparameter Tuning Process: African Vulture Optimization Algorithm
Performance Measures: Recall Rate, Precision Rate, Accuracy Rate, F-score Rate

FIGURE 1. The overall procedure of the BDEV-CAML system.

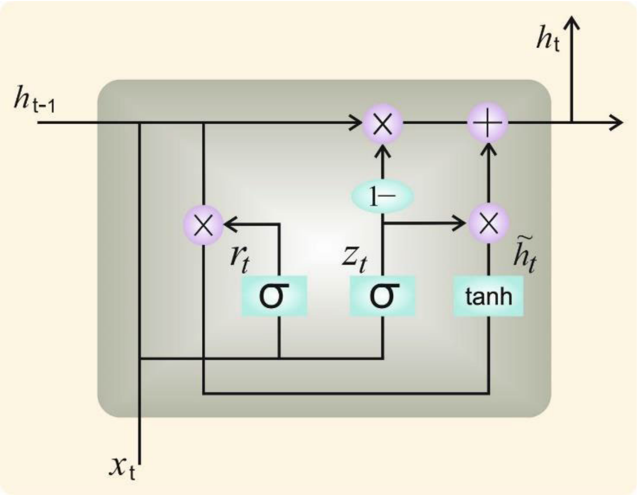


FIGURE 2. The architecture of GRU.

models accomplish decreasing PR of 98.6%, 97.06%, and 95.87% correspondingly. Simultaneously, with IOHog fault class, the BDEV-CAML system obtains increasing PR of 99.72% while the PSO-DAWRF, DAWRF, and RF methods achieve decreasing PR of 97.91%, 96.41%, and 95.25% respectively.

In Table 2, and Fig. 4, a comparative recall rate (RR) analysis of the BDEV-CAML method is given. The results show that the RF method accomplishes minimum efficiency, whereas the PSO-DAWRF and DAWRF methods achieve closer outcomes. Nonetheless, the BDEV-CAML method obtains increasing RR values. For example, with CPUHog fault class, the BDEV-CAML technique reaches a better RR of 99.49% while the PSO-DAWRF, DAWRF, and RF methods achieve a reducing RR of 98.51%, 97.55%, and 96.32% correspondingly. Simultaneously, with IOHog fault class, the BDEV-CAML method accomplished maximum RR of 99.61% while the PSO-DAWRF, DAWRF, and RF methods attained minimum RR of 98.06%, 96.86%, and 95.51% correspondingly.

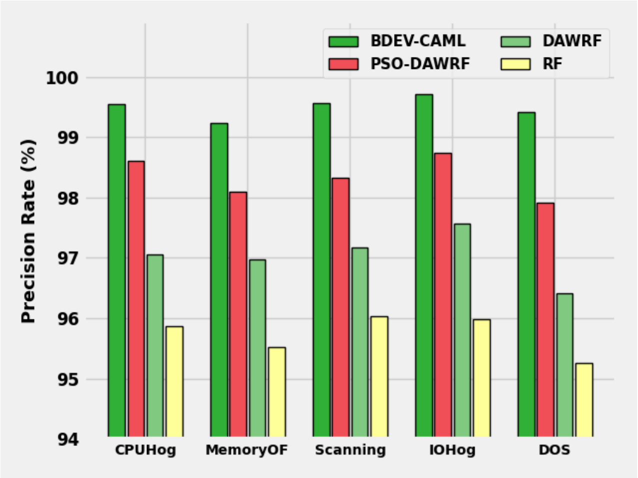


FIGURE 3. PR analysis of BDEV-CAML approach under varying classes.

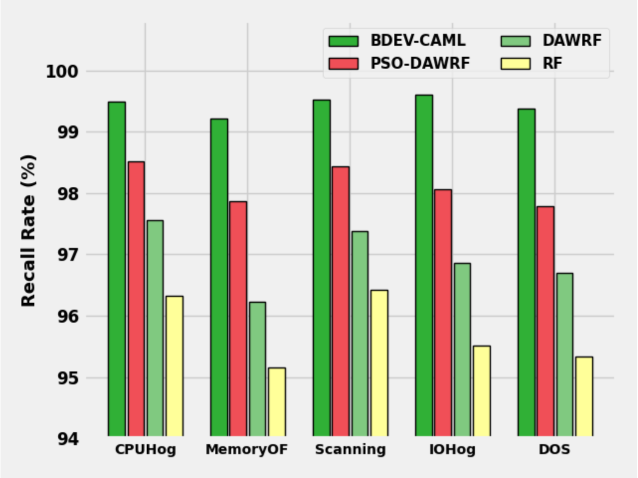


FIGURE 4. RR analysis of BDEV-CAML approach under varying classes.

In Table 3, and Fig. 5, a comparative accuracy rate (AR) analysis of the BDEV-CAML method is given. The results show that the RF method accomplishes the least efficiency, whereas the PSO-DAWRF and DAWRF methods attain closer results. Nonetheless, the BDEV-CAML method reaches better AR values. For the case with CPUHog fault class, the BDEV-CAML method obtains an increasing AR of 99.62% while the PSO-DAWRF, DAWRF, and RF methods achieve a minimum AR of 98.44%, 96.78%, and 95.07% correspondingly. Simultaneously, with IOHog fault class, the BDEV-CAML method obtains a maximum AR of 99.53%, while the PSO-DAWRF, DAWRF, and RF methods achieve a

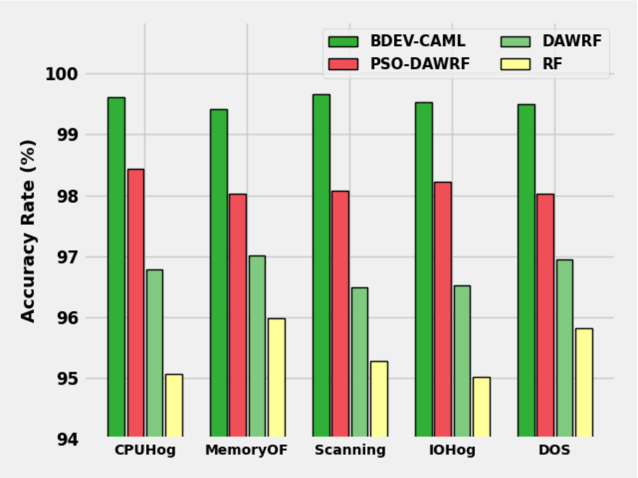


FIGURE 5. AR analysis of BDEV-CAML approach under varying classes.

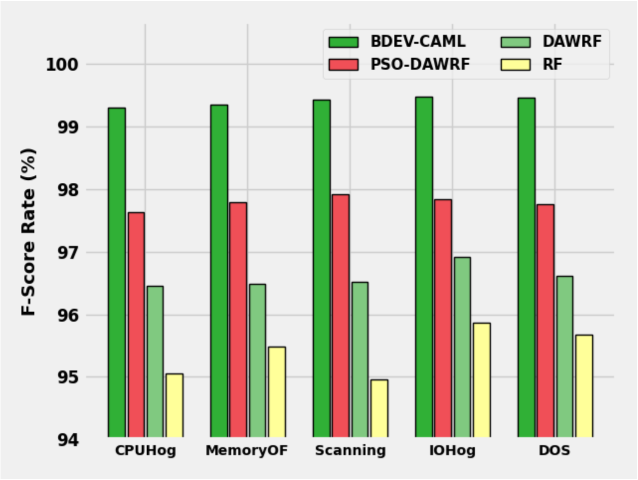


FIGURE 6. FR analysis of BDEV-CAML approach under varying classes.

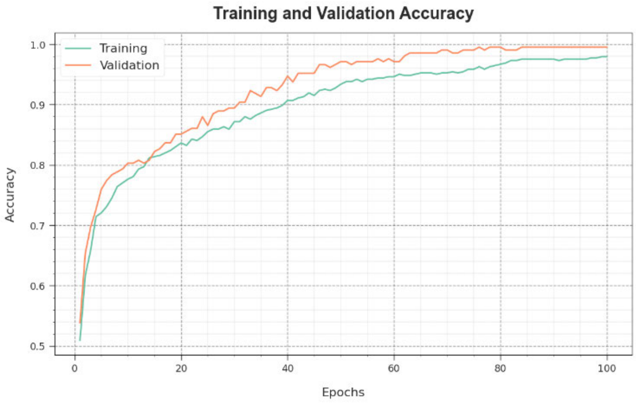


FIGURE 7. Accuracy curve of the BDEV-CAML approach.

minimum AR of 98.03%, 96.94%, and 95.81% correspondingly.

In Table 4, and Fig. 6, a comparative F-score rate (FR) analysis of the BDEV-CAML method is given. The results indicate that the RF model reaches the tiniest efficiency, whereas the PSO-DAWRF and DAWRF methods attain closer results. Nonetheless, the BDEV-CAML technique obtains better FR values. For example, with CPUHog fault class, the BDEV-CAML technique gets maximum FR of 99.55% while the PSO-DAWRF, DAWRF, and RF models

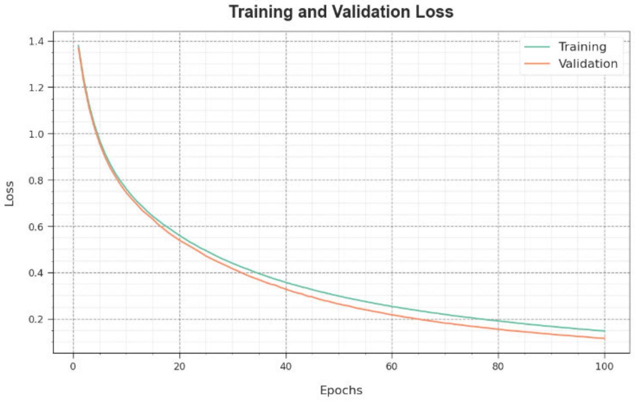


FIGURE 8. Loss curve of the BDEV-CAML approach.

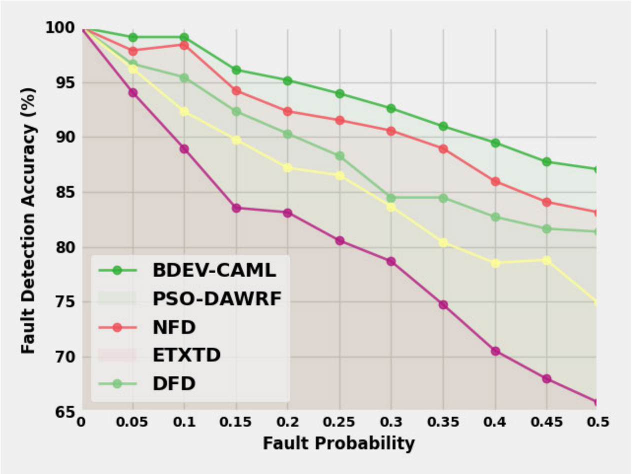


FIGURE 9. FDA analysis of BDEV-CAML approach with other existing methods.

TABLE 2. Recall rate analysis of BDEV-CAML approach with other systems under varying classes.

Recall Rate (%)				
Class	BDEV-CAML	PSO-DAWRF	DAWRF	RF
CPUHog	99.49	98.51	97.55	96.32
MemoryOF	99.22	97.86	96.22	95.16
Scanning	99.53	98.43	97.38	96.42
IOHog	99.61	98.06	96.86	95.51
DOS	99.38	97.79	96.70	95.34

TABLE 3. Accuracy rate analysis of the BDEV-CAML approach with other systems under varying classes.

Accuracy Rate (%)				
Class	BDEV-CAML	PSO-DAWRF	DAWRF	RF
CPUHog	99.62	98.44	96.78	95.07
MemoryOF	99.42	98.02	97.01	95.98
Scanning	99.66	98.07	96.49	95.27
IOHog	99.53	98.22	96.52	95.01
DOS	99.50	98.03	96.94	95.81

TABLE 4. F-score rate analysis of BDEV-CAML approach with other systems under varying classes.

F-Score Rate (%)				
Class	BDEV-CAML	PSO-DAWRF	DAWRF	RF
CPUHog	99.30	97.64	96.46	95.05
MemoryOF	99.36	97.79	96.49	95.49
Scanning	99.43	97.92	96.52	94.96
IOHog	99.48	97.84	96.91	95.86
DOS	99.46	97.76	96.61	95.67

accomplish minimum FR of 98.6%, 97.06%, and 95.87% correspondingly.

Simultaneously, with IOHog fault class, the BDEV-CAML method accomplishes a maximum FR of 99.72%, while the PSO-DAWRF, DAWRF, and RF methods obtain a minimum FR of 97.91%, 96.41%, and 95.25% correspondingly.

Fig. 7 examines the accuracy of the BDEV-CAML method during the training and validation process on the test dataset. The figure indicates that the BDEV-CAML method accomplishes increasing accuracy values over epochs. In addition, the increasing validation accuracy over training accuracy shows that the BDEV-CAML method learns efficiently on the test dataset.

The loss analysis of the BDEV-CAML technique at the time of training and validation is demonstrated on the test dataset in Fig. 8. The results show that the BDEV-CAML technique obtains closer training and validation loss values. Note that the BDEV-CAML method learns efficiently on the test dataset.

Table 5 and Fig. 9 demonstrate the fault detection accuracy (FDA) results of the BDEV-CAML technique with recent models [12].

TABLE 5. FDA analysis of BDEV-CAML approach with other existing methods [12].

Fault Detection Accuracy (%)					
Fault Probability	BDEV-CAML	PSO-DAWRF	NFD	ETXTD	DFD
0.00	100.00	100.00	100.00	100.00	100.00
0.05	99.10	97.88	96.67	96.26	94.10
0.10	99.10	98.42	95.45	92.34	88.96
0.15	96.12	94.23	92.34	89.77	83.55
0.20	95.18	92.34	90.31	87.20	83.14
0.25	93.96	91.53	88.28	86.52	80.57
0.30	92.61	90.58	84.49	83.68	78.68
0.35	90.98	88.96	84.49	80.43	74.75
0.40	89.50	85.98	82.73	78.54	70.56
0.45	87.74	84.09	81.65	78.81	67.99
0.50	87.06	83.14	81.38	74.89	65.83

The results show better outcomes of the BDEV-CAML technique with increasing FDA values under all values of fault probability (FP). For instance, with FP of 0.05, the BDEV-CAML technique gains increasing FDA of 99.10% while the PSO-DAWRF, NFD, ETXTD, and DFD models obtain reducing FDA of 97.88%, 96.67%, 96.26%, and 94.10% respectively. Simultaneously, with FP of 0.50, the BDEV-CAML technique gains an increasing FDA of 87.06% while the PSO-DAWRF, NFD, ETXTD, and DFD methods acquire minimum FDA 83.14%, 81.38%, 74.89%, and 65.83% correspondingly. These results assured the improved performance of the BDEV-CAML technique over other existing models.

V. CONCLUSION

In this study, we have designed a novel BDEV-CAML algorithm for the effectual identification of faults in the IoT environment. The presented BDEV-CAML technique integrated the advantages of the blockchain, IoT, and ML concepts for boosting the trustworthiness, efficacy, and security of the IoT network. In BC technology, IoT devices with highly decentralized decision-making capability can attain a consensus on the efficiency of intrablock transactions. For fault detection in the IoT network, the AVOA with DBiGRU technique is utilized for the optimal hyperparameter tuning of the DBiGRU model, which helps in improving the fault detection rate. A detailed set of experiments were carried out to highlight the enhanced performance of the BDEV-CAML algorithm. The comprehensive experimental results stated the improved performance of the BDEV-CAML technique over other existing models. In future, hybrid DL models can boost the performance of the BDEV-CAML technique. In addition, the proposed model can be extended to the detection of the faults in the real time IoT environment.

REFERENCES

[1] W. Zhang, Q. Lu, Q. Yu, Z. Li, Y. Liu, S. K. Lo, S. Chen, X. Xu, and L. Zhu, "Blockchain-based federated learning for device failure detection in industrial IoT," *IEEE Internet Things J.*, vol. 8, no. 7, pp. 5926–5937, Apr. 2021.

[2] W. Zhang, Z. Wang, and X. Li, "Blockchain-based decentralized federated transfer learning methodology for collaborative machinery fault diagnosis," *Rel. Eng. Syst. Saf.*, vol. 229, Jan. 2023, Art. no. 108885.

[3] J. Liu, H. Liu, C. Chakraborty, K. Yu, X. Shao, and Z. Ma, "Cascade learning embedded vision inspection of rail fastener by using a fault detection IoT vehicle," *IEEE Internet Things J.*, vol. 10, no. 4, pp. 3006–3017, Feb. 2023.

[4] T. Mezair, Y. Djenouri, A. Belhadi, G. Srivastava, and J. C.-W. Lin, "A sustainable deep learning framework for fault detection in 6G industry 4.0 heterogeneous data environments," *Comput. Commun.*, vol. 187, pp. 164–171, Apr. 2022.

[5] M. Ul Mehmood, A. Ulasyar, A. Khattak, K. Imran, H. S. Zad, and S. Nisar, "Cloud based IoT solution for fault detection and localization in power distribution systems," *Energies*, vol. 13, no. 11, p. 2686, May 2020.

[6] S. U. Jan, Y. D. Lee, and I. S. Koo, "A distributed sensor-fault detection and diagnosis framework using machine learning," *Inf. Sci.*, vol. 547, pp. 777–796, Feb. 2021.

[7] H. Zhang, R. Li, and C. Shi, "Deep learning technology of Internet of Things blockchain in distribution network faults," *J. Intell. Syst.*, vol. 31, no. 1, pp. 965–978, Aug. 2022.

[8] A. H. Sodhro, A. Lakhan, S. Pirbhulal, T. M. Groenli, and H. Abie, "A lightweight security scheme for failure detection in microservices IoT-edge networks," in *Sensing Technology*. Cham, Switzerland: Springer 2022, pp. 397–409.

[9] Y. Jiahao, X. Jiang, S. Wang, K. Jiang, and X. Yu, "SVM-BiLSTM: A fault detection method for the gas station IoT system based on deep learning," *IEEE Access*, vol. 8, pp. 203712–203723, 2020.

[10] M. Dzaferagic, N. Marchetti, and I. Macaluso, "Fault detection and classification in industrial IoT in case of missing sensor data," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 8892–8900, Jun. 2022.

[11] M. Trivedi, R. Kakkar, R. Gupta, S. Agrawal, S. Tanwar, V.-C. Niculescu, M. S. Raboaca, F. Alqahtani, A. Saad, and A. Tolba, "Blockchain and deep learning-based fault detection framework for electric vehicles," *Mathematics*, vol. 10, no. 19, p. 3626, Oct. 2022.

[12] W. Zhang, J. Wang, G. Han, S. Huang, Y. Feng, and L. Shu, "A data set accuracy weighted random forest algorithm for IoT fault detection based on edge computing and blockchain," *IEEE Internet Things J.*, vol. 8, no. 4, pp. 2354–2363, Feb. 2021.

[13] A. Belhadi, Y. Djenouri, G. Srivastava, A. Jolfaei, and J. C.-W. Lin, "Privacy reinforcement learning for faults detection in the smart grid," *Ad Hoc Netw.*, vol. 119, Aug. 2021, Art. no. 102541.

[14] H. Huang, S. Ding, L. Zhao, H. Huang, L. Chen, H. Gao, and S. H. Ahmed, "Real-time fault detection for IIoT facilities using GBRBM-based DNN," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 5713–5722, Jul. 2020.

[15] S. Lavanya, A. Prasanth, S. Jayachitra, and A. Shenbagarajan, "A tuned classification approach for efficient heterogeneous fault diagnosis in IoT-enabled WSN applications," *Measurement*, vol. 183, Oct. 2021, Art. no. 109771.

[16] S. Mittra, A. Aprameya, and B. K. Mohanta, "Smart grid power theft and fault detection using IoT and blockchain," in *Proc. Int. Conf. Advancements Electr., Electron., Commun., Comput. Autom. (ICAECA)*, Oct. 2021, pp. 1–5.

[17] Z. Zhang, C. Guan, H. Chen, X. Yang, W. Gong, and A. Yang, "Adaptive privacy-preserving federated learning for fault diagnosis in Internet of Ships," *IEEE Internet Things J.*, vol. 9, no. 9, pp. 6844–6854, May 2022.

[18] L. Xia, Y. Liang, J. Leng, and P. Zheng, "Maintenance planning recommendation of complex industrial equipment based on knowledge graph and graph neural network," *Rel. Eng. Syst. Saf.*, vol. 232, Apr. 2023, Art. no. 109068.

[19] L. Xia, P. Zheng, X. Huang, and C. Liu, "A novel hypergraph convolution network-based approach for predicting the material removal rate in chemical mechanical planarization," *J. Intell. Manuf.*, vol. 33, no. 8, pp. 2295–2306, Dec. 2022.

[20] L. Xia, P. Zheng, J. Li, W. Tang, and X. Zhang, "Privacy-preserving gradient boosting tree: Vertical federated learning for collaborative bearing fault diagnosis," *IET Collaborative Intell. Manuf.*, vol. 4, no. 3, pp. 208–219, Sep. 2022.

[21] L. Xia, Y. Liang, P. Zheng, and X. Huang, "Residual-hypergraph convolution network: A model-based and data-driven integrated approach for fault diagnosis in complex equipment," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023.

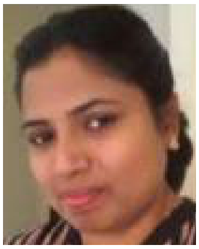
[22] K. L. Keung, Y. Y. Chan, K. K. H. Ng, S. L. Mak, C. H. Li, Y. Qin, and C. W. Yu, "Edge intelligence and agnostic robotic paradigm in resource synchronisation and sharing in flexible robotic and facility control system," *Adv. Eng. Informat.*, vol. 52, Apr. 2022, Art. no. 101530.

[23] K. L. Keung, C. K. M. Lee, and P. Ji, "Industrial Internet of Things-driven storage location assignment and order picking in a resource synchronization and sharing-based robotic mobile fulfillment system," *Adv. Eng. Informat.*, vol. 52, Apr. 2022, Art. no. 101540.

[24] E. A. Adeniyi, R. O. Ogundokun, S. Misra, J. B. Awotunde, and K. M. Abiodun, "Enhanced security and privacy issues in a multi-tenant environment of green computing using blockchain technology," in *Blockchain Applications in the Smart Era*. Cham, Switzerland: Springer, 2022, pp. 65–83.

[25] M. Lu and Q. Xie, "A novel approach for spatially controllable high-frequency forecasts of park visitation integrating attention-based deep learning methods and location-based services," *ISPRS Int. J. Geo-Information*, vol. 12, no. 3, p. 98, Feb. 2023.

[26] S. R. Nayak, R. K. Khadanga, S. Panda, P. R. Sahu, S. Padhy, and T. S. Ustun, "Participation of renewable energy sources in the frequency regulation issues of a five-area hybrid power system utilizing a sine cosine-adopted African vulture optimization algorithm," *Energies*, vol. 16, no. 2, p. 926, Jan. 2023.



THAVAVEL VAIYAPURI (Member, IEEE) is currently an Assistant Professor with the College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University. With nearly 20 years of research and teaching experience, she has published more than 50 research papers in impacted journals and international conferences. Her research interests include data science, security, computer vision, and high-performance computing. She is a member of the IEEE Computer Society. She is a fellow of HEA, U.K.

K. SHANKAR (Senior Member, IEEE) received the Ph.D. degree in computer science from Alagappa University, Karaikudi, India. He is currently a Postdoctoral Fellow with the Big Data and Machine Learning Laboratory, South Ural State University, Russia, and an Adjunct Faculty of the Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India. He has authored/coauthored over 150 ISI journal articles (with a total impact factor of more than 350) and more than 100 Scopus-indexed articles. Besides, he has authored/edited six international books published by recognized publishers, such as Springer and CRC. His current research interests include healthcare applications, secret image sharing schemes, digital image security, cryptography, the Internet of Things, and optimization algorithms.

SURENDRAN RAJENDRAN (Member, IEEE) received the Ph.D. degree in computer science and engineering from Sathyabama University, in 2014. He is currently a Professor with the Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, India. His research interests include cloud computing, the Internet of Things, deep learning, artificial intelligence, and big data.

SACHIN KUMAR (Senior Member, IEEE) received the Ph.D. degree in data mining/machine learning from the Indian Institute of Technology Roorkee, in 2017. He is currently with the Department of Computer Science, South Ural State University, Chelyabinsk, Russia. He is also the Head of the Data Mining and Virtualization Laboratory and a leading Researcher with the Big Data and Machine Learning Research Laboratory. His research interests include intelligent transportation systems, machine learning, data mining, the IoT, and health informatics. He also served as a reviewer for various reputed international journals.

SRIJANA ACHARYA received the B.S. degree in computer application from MCRP University, India, in 2003, and the M.S. degree in information and communication engineering and the Ph.D. degree in digital convergence business from Yeungnam University, in 2014 and 2021, respectively. She received various scholarships for her M.S. and Ph.D. studies. She is currently a Postdoctoral Researcher with the Department of Convergence Science, Kongju National University, South Korea. Her research interests include webometrics, open data, data security, SNS security, SNS analysis, knowledge management, and digital convergence.

HYUNIL KIM received the B.S. degree in applied mathematics and the M.S. and Ph.D. degrees in information security from Kongju National University, South Korea, in 2014, 2016, and 2019, respectively. Also, he was a Postdoctoral Researcher at the Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu, South Korea. He is currently a Research Professor with Kongju National University. His research interests include AI security with federated learning, decentralized identifiers, and blockchain.