IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 29 August 2023, accepted 21 September 2023, date of publication 29 September 2023, date of current version 10 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3320716

RESEARCH ARTICLE

Collaborative Worker Safety Prediction Mechanism Using Federated Learning Assisted Edge Intelligence in Outdoor Construction Environment

SA JIM SOE MOE[1], BONG WAN KIM[2], ANAM NAWAZ KHAN[1], XU RONGXU[3], NGUYEN ANH TUAN[1], KWANGSOO KIM[2], AND DO HYEUN KIM[1]

[1]Department of Computer Engineering, Jeju National University, Jeju-si 63243, Republic of Korea
[2]Electronics and Telecommunications Research Institute (ETRI), Daejeon 34129, Republic of Korea
[3]Big Data Research Center, Jeju National University, Jeju 63243, Republic of Korea

Corresponding author: Do Hyeun Kim (kimdh@jejunu.ac.kr)

This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 22SMIP-A156371-03). Any correspondence related to this paper should be addressed to Do Hyeun Kim.

ABSTRACT Monitoring construction site safety through physical observations is inherently flawed due to the complex and dynamic nature of construction sites. To overcome these challenges and enhance worker safety management, decentralized model training-assisted edge intelligence emerges as a promising solution. However, despite the potential benefits, our investigation reveals that no research for worker safety prediction has been grounded in the Federated Learning (FL) approach. In this context, we present a novel approach to worker safety prediction, leveraging FL in outdoor construction environments while preserving the privacy and security of sensitive data. Our methodology involves deploying sensor-based IoT devices at construction sites to collect highly granular spatial and temporal weather, building, and worker data. This data is then collaboratively utilized for training Deep Neural Network (DNN) models on the edge nodes in a cross-silos manner. To implement our approach, we establish a test-bed utilizing the EdgeX framework and constrained devices such as Raspberry Pi 4Bs, acting as edge nodes. Following the collaborative training, the resultant global model is deployed on participating nodes for edge inference, ensuring optimal network resource utilization and data privacy. The experimental results demonstrate the efficacy of the proposed approach in improving the utilization of construction safety management systems and reducing the risk of accidents and fatalities in the future. The outcome is a system that exhibits enhanced speed and responsiveness, a crucial aspect for time-sensitive applications such as the prediction of worker safety.

INDEX TERMS Worker safety, outdoor construction site, federated learning, edge computing, EdgeX, Internet of Things.

I. INTRODUCTION

The world has undergone rapid change in recent decades, and the construction industry has been at the forefront of this transformation. With an urban population that grows by 200,000 people every day, it is clear that the global construction industry has been greatly influenced by these demographic shifts [1]. However, due to its dynamic, constantly shifting, and heterogeneous spatiotemporal environment, construction is considered one of the most dangerous industries for workers. Worker safety is an ongoing issue that requires continued attention and effort. A recent study suggests [2] that workers frequently encounter potential safety and health threats during the building process as a result of the hazardous working conditions at construction sites. According to the information presented in [3], it is estimated that

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

a significant number, possibly exceeding 6,500, of migrant workers have tragically lost their lives on construction sites in Qatar, as the country prepares for the 2022 World Cup through major projects. Therefore, an intelligent solution based on predictive analytics is required to analyze the data of the construction industry to cope with the challenges faced by the construction industry and enhance the safety paradigm for workers.

The Internet of Things (IoT) is a next-generation technology that aims to bridge the physical and digital worlds. It allows for the interconnection of computing devices that are embedded in everyday objects, enabling them to transmit data via the Internet [4]. It is a network system that utilizes both wired and wireless technologies to connect physical devices and the sensors embedded within or attached to them to the Internet. The IoT encompasses a wide range of electronic devices that serve as signal converters, transforming real-world data into digital or analog formats that can be understood by both humans and machines. These devices can take input from various sources and include nodes such as Bluetooth, Wi-Fi, and ZigBee, which can be installed in vehicles or homes [5]. With its flexibility, IoT technology can be used for a variety of applications and use cases. Current research efforts are focused on reducing construction site accidents by integrating safety measures and information and communication technology. For instance, the study in [6] devised an IoT-based system for monitoring indoor safety in the context of COVID-19. Additionally, the authors of the paper [7] prominently utilized IoT by integrating IoT sensors into a real-time surveillance helmet. These sensors assess the immediate surroundings of the individual, monitor their health in real-time, and initiate appropriate actions when the working conditions become unfavorable for the miner.

Due to the availability of the massive amount of data generated by IoT devices, it is quite possible to employ data-driven assisted predictive analytic techniques to facilitate safety management for formulating efficient decisions to improve safety at the workplace. Predictive analytics [8] is a field of data analysis that employs statistical models, machine learning algorithms, and data mining techniques to recognize patterns and anticipate future events and trends. This approach leverages historical data to inform predictions and is widely adopted in forecasting [9], risk assessment [10], fraud detection [11], marketing optimization [12], etc. The ultimate objective is to harness the power of data to unlock valuable insights and make data-driven decisions that drive business growth and success. In this study, this technique is employed to infer worker safety by leveraging data from IoT devices deployed in the construction environment.

Integrating predictive analytics techniques with data generated from IoT devices presents a challenge in handling vast amounts of data. Therefore, machine learning is necessary to extract and process useful knowledge and underlying patterns from the immense data and construct an effective predictive model. Machine learning has proven to be a valuable tool, especially in handling massive amounts of data, providing significant time savings, and maximizing computing resources. Predictive modeling techniques are integral components of machine learning-based systems that enable organizations to anticipate future demands using inferred knowledge, leading to better-informed decision-making. Thanks to advancements in machine learning models, it has become increasingly feasible to create reliable and efficient prediction models that leverage discovered knowledge to accurately forecast future outcomes. There exist various machine learning algorithms that can be employed to analyze the identified insights from past data. Some of the commonly used algorithms are Linear Regression [13], Logistic Regression [14], Decision Trees [15], Support Vector Machines (SVM) [16], Random Forests [17], and so on. Moreover, an updated machine learning technique, such as Automated Machine Learning (AutoML), has enhanced its workflow by automating and optimizing substantial expertise in various areas, including data preprocessing, feature engineering, model selection, and hyperparameter tuning. In a related study [18], the authors leveraged this advanced machine-learning technique to predict carpark price indices. Yet, our study has chosen to adopt a manual approach utilizing DNN, a deep learning model.

DNN has become a prominent algorithm in the field of deep learning research, enabling the development of intelligent applications in diverse domains [19]. DNN is a form of Artificial Neural Network (ANN) that consists of multiple layers of interlinked nodes or neurons, which process input data and progressively extract increasingly abstract features from the input. The deep structure of DNN allows them to learn hierarchical representations of complex patterns and relationships in the data, making them well-suited for tasks such as image and speech recognition [20], natural language processing [21], and predictive modeling [22]. DNN employs backpropagation for weight adjustment between neurons, minimizing the difference between predicted outputs and target values. This supervised learning technique automates the feature representation process, eliminating the need for manual feature engineering traditionally used in other models. DNN is renowned for its effectiveness in the realm of deep learning [23].

However, traditional machine learning often uses a centralized learning approach where data from various sources is collected at one location, and the model is trained on that data before being sent back to the sources. Therefore, this approach has several drawbacks, including high communication and computation costs, limited adaptability, and privacy concerns.

On the other hand, FL, introduced by McMahan et al. [24] in 2017, is a decentralized approach where data remains on individual devices, and the model is trained locally on each device. To enable collaborative learning on a large scale within FL, the Federated Averaging (FedAVG) technique comes into play. FedAVG leverages the collective knowledge

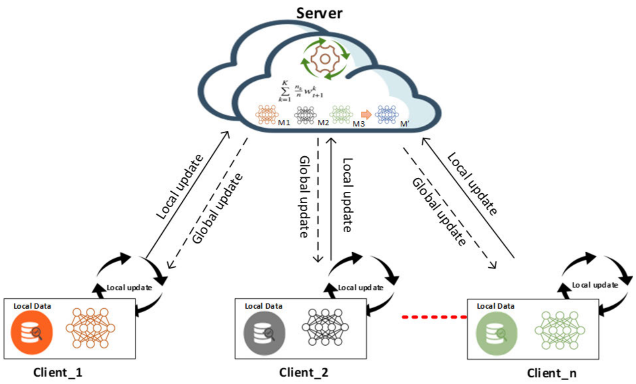


FIGURE 1. Federated learning process [24].

of distributed devices while ensuring data privacy. In this process, the model updates generated on each device are sent to a central server. The server then aggregates these updates, as illustrated in Figure 1, to create a global model that benefits from the combined intelligence of all participating devices.

Worker safety prediction is a crucial concern in outdoor construction environments, given the numerous challenges posed by factors like fluctuating weather conditions, building conditions, and worker well-being. Surprisingly, previous research has not specifically explored the use of FL to tackle this issue. Moreover, the study in [25] emphasized the significance of deep learning as an emerging field within construction safety, underscoring the ongoing need to maximize the potential of AI technology. Consequently, our research aims to bridge this gap by employing FL to train a machine-learning model capable of effectively predicting worker safety in the outdoor construction environment. The main contributions of this paper are as follows:

- Development of an intelligent worker safety edge inference system based on the FL for effective construction safety management to proactively prevent potential accidents and fatalities.
- Deployment of collaboratively trained model on edge nodes to infer worker safety in real-time under dynamically changing data collected from various sensors including weather conditions, building circumstances, and worker status data.
- Development of an intelligent edge analytic framework using EdgeX to provide multiple service scenarios such as device and data management as well as interoperability between IoT devices and safety service applications at the network edge.
- Design and development of a user-friendly web application that enables users to interact with the system through a user interface (UI) for easy management, operation, and assessment of worker safety.
- Provide extensive experimental results to evaluate the effectiveness of the proposed FL-based technique for ensuring the reliability of this study to enhance construction safety management.

The following structure is adopted in the remainder of this paper: Section II will discuss the related works, followed by Section III which presents the proposed approach. The implementation environment and corresponding results will be presented in Section IV, while Section V will provide an in-depth explanation of the performance analysis. Finally, section VI concludes the paper and presents our future works.

II. LITERATURE REVIEW

The construction industry is plagued by a pressing issue of safety, which poses a significant threat to the well-being of its workers. Inappropriate working conditions and frequent workplace accidents are unfortunately common, causing construction employees to miss work and impacting their health and livelihood. Hence, worker safety prediction is essential in construction sites because it enables us to anticipate and mitigate potential hazards or threats before they occur. By predicting possible risks or safety issues, we can take proactive measures to prevent or minimize their impact, ultimately preventing injuries, damages, or even fatalities.

A. OUTDOOR CONSTRUCTION ENVIRONMENT BASED ON IOT

Adopting innovative technologies in construction has the potential to bring numerous benefits, with enhanced safety being of paramount importance. By ensuring the well-being of workers and the smooth operation of construction sites, these advancements can play a key role in securing the success of projects and the overall growth of the industry [26], [27]. The rapid adoption of IoT and other cutting-edge technologies has brought about heightened concerns for safety on construction sites, as their integration into the industry has uncovered new risks and challenges [28]. According to the findings in [29], IoT is utilized as a tool for sharing construction safety knowledge. Also, the authors of [30], introduced a system that employs IoT sensors to monitor safety hooks, designed to prevent falls from heights. This system effectively automates real-time safety monitoring for numerous workers at complex construction sites, providing valuable support to safety managers. Integrating sensors with construction workers can be highly advantageous in implementing IoT for safety objectives. Numerous models and prospective approaches are available for the integration of IoT sensors with protective gear such as helmets and clothing, offering a more comprehensive and effective solution for safeguarding workers on construction sites [31]. To effectively identify various hazards present on construction sites, different types of sensors are utilized, including infrared, accelerometer, RFID, and gyroscope sensors, which are among the most widely used. The application of these tools is vital in enhancing safety and reducing the risks inherent in construction work [32]. Additionally, monitoring the surrounding site conditions is possible using GPS to identify potentially hazardous areas on the building site [33] and prevent collisions through the use of ultrasound sensors [34]. This helps ensure a safe working environment and reduces the risk of accidents.

B. CENTRALIZED TRAINING APPROACHES FOR WORKER SAFETY PREDICTION

Predicting safety or risk is crucial as it enables us to anticipate and mitigate potential hazards or threats before they happen. Machine learning and AI-based approaches are extensively being applied to safety prediction. For instance, the authors in [35] developed a model using Random Forest (RF) to utilize feature importance analysis to identify the key factors that contribute to occupational accidents at construction sites and develop a model for classifying and predicting such accidents. They also examined the correlation between these factors and types of occupational accidents. However, their research needs to be further improved in the future, such as collecting real-time data and developing models that can accurately predict occupational accident types. In [36], the authors present a hybrid method that combines an Adaptive Neural Network-based Fuzzy Inference System (ANFIS) and a safety inspection checklist to detect risk factors and forecast the likelihood of scaffold falls at construction sites. Their approach facilitates the identification and evaluation of critical conditions and scenarios that have a substantial impact on the risk of falling. However, further research is needed to investigate the variables that influence the safety of other types of scaffolding. Choi et al. [37] has developed a highly effective prediction model that can pinpoint the probable danger of fatal accidents in construction sites. The machine learning algorithms used in the model rely on the industrial accident data collected between 2011 and 2016 by the Ministry of Employment and Labor (MOEL) in the Republic of Korea. Their model's performance evaluation showed an exceptional predictive rate for accurately classifying workers who might be facing a fatality risk. Yet, they stated the use of historical data is limited because of the privacy information law in the country. For observing and assessing how workers and equipment interact, [38] proposed a methodology that employs computer vision and deep learning techniques. This methodology detects the locations and trajectories of workers and equipment and identifies danger zones. However, they also mentioned that the detection accuracy can be further improved by enriching the dataset as the training dataset is limited. To resolve the problems and challenges associated with centralized training schemes, FL, [24] a promising approach is utilized.

C. FEDERATED LEARNING FOR HEALTH AND SAFETY

FL has emerged as an effective solution to address user privacy concerns, thereby enabling the collection of larger datasets to train machine learning models and enhance their accuracy and efficiency. With its demonstrated potential across a range of industries and engineering disciplines, it's no surprise that FL has become a rapidly growing area of research, with numerous successful outcomes. Recently, many researchers have focused on developing decentralized solutions for ensuring human health and safety. The authors of [39] leveraged the power of FL, a privacy-preserving machine learning method, to predict 7-day mortality rates among hospitalized COVID-19 patients. Their use of FL allowed them to build robust predictive models without the need to centrally collect raw clinical data from multiple institutions. The results of their study demonstrate the potential of FL in healthcare. Borger et al. [40] conducted research on the application of FL in clinical Natural Language Processing. Their study focused on using this approach to assess the risk of violence in a cross-institutional psychiatric setting. The results of their investigation demonstrated that FL can be effectively employed in such a setting and could lead to the development of novel applications for this approach based on clinical notes. Li et al. [41] proposed a framework called FedSWP, which employs federated transfer learning to enable Smart Work Packaging (SWP) and safeguard the private image data of construction workers in Occupational Health and Safety (OHS) management. Their research serves as a stepping stone towards expanding and adapting FedSWP for various OHS applications in the construction industry. Qi et al. [42] have introduced a powerful deep learning framework, named FedAGCN, that utilizes both FL and asynchronous graph convolutional networks to accurately predict real-time traffic flow. Additionally, they have proposed a new graph FL strategy, called GraphFed, which effectively reduces the time required to train the deep learning model. Therefore, FL has demonstrated its ability to overcome the limitations of centralized machine learning, not only in addressing privacy concerns but also in enhancing model efficiency through the training of models using data from diverse entities that are usually isolated in data silos, and by integrating the knowledge gained into a globally trained model. However, till now, FL has not been used for worker safety prediction in the outdoor construction environment. To the best of our knowledge, the proposed study is the first attempt to develop an intelligent edge inference system established upon FL for proactive worker safety that is accurate even under dynamic changing conditions.

III. PROPOSED METHODOLOGY

The proposed mechanism for collaborative worker safety prediction is deployed at the Edge node network, harnessing the capabilities of a renowned distributed machine learning technique called federated learning. Specifically, we employ the FedAVG technique, which facilitates the averaging of weight updates from all participating clients. Moreover, this approach involves leveraging data generated by IoT devices. Multiple edge nodes simulate the contribution of their sensors, representing various construction workers. By incorporating data from diverse sources such as weather information, building conditions, and worker status, this study aims to enhance worker safety prediction in outdoor construction environments. To achieve this objective, the DNN model serves as the foundational learner, featuring

8 input dimensions, two densely connected hidden layers, and utilizing the ReLU activation function. Normalization is performed individually on each node's data to handle the varying magnitudes of readings from sensors.

Figure 2 presents the conceptual diagram of this study, which predicts worker safety in the outdoor construction environment. The mechanism involves multiple components, including an edge server, edge nodes, IoT devices, and clients. To implement this, three phases of operation are conducted, including the initialization phase, the operation phase, and the inference phase. Particularly, multiple clients (also known as edge nodes) are registered. Subsequently, a task, which is setting experimental configurations of FL, is generated during the mechanism's initialization phase. The edge server then interacts with the edge nodes through FL to acquire the optimal global model for predicting worker safety during the training phase. FL allows for the collaborative training of a single DNN model across multiple edge nodes, without requiring these nodes to send their own data to a central aggregation server. To accomplish this goal, FedAVG which enables the edge devices to perform multiple iterations of weight updates to refine their local models before transmitting the updated weights to the central server is utilized. This involves independent training of the DNN model across the clients on their local data, with the parameters being averaged periodically on the server. Detailed information on the process of FL and the FedAVG algorithm utilized in our approach, is provided in subsections III-A and III-B. Once the optimal global model is developed as a result of collaborative training, it is deployed to registered nodes for inferring worker safety. IoT devices that collect data from the outdoor construction environment are connected to the micro-service module of EdgeX. Based on the inference results the developed system later provides the desired actuation by triggering the control commands. EdgeX microservices help achieve better device management, and connectivity, thus allowing IoT devices to be easily controlled and monitored. During the inference phase, IoT devices' data are derived from the EdgeX repository and preprocessed on the edge nodes. Then preprocessed data is fed into the prediction model and inference results are achieved. Furthermore, users as well as clients can provision, manage, and operate the system through the web service provided by the edge server, which can be accessed via the internet.

Figure 3 shows the layered architecture of our mechanism which consists of three layers: the client layer, the edge layer, and the device layer. For gathering information from the outdoor construction setting, the device layer of our mechanism includes sensors that are mounted to the IoT devices. These WiFi-enabled devices continuously collect data from their sensors and transmit it to the internet over HTTP. In our mechanism, devices are represented by their resources and are registered and managed on the edge node, where data is collected through edge computing services, which are microservices of EdgeX. The network's available resources enable sensing and actuating functionalities, which in turn provide IoT services to the edge node. IoT frameworks and libraries are utilized to create these resources, which empower IoT devices to provide services by transmitting data. Event management is employed to publish sensing data upon request from the edge node.

The edge layer of the mechanism consists of the edge server and edge nodes, which work together to derive the optimal global model through FL using historical data at the edge of the network. Once the optimal global model is determined, it is first deployed on the node and is used to predict the worker safety of new data. When a client requests to predict worker safety, data is first collected from the IoT devices and transferred to the repository of EdgeX. The safety inference result provided by the edge node is then sent back to the edge server to display to the client through UI. Besides this, the client service provider facilitates access to virtual objects for users within the system and connects with physical devices over the Internet.

The client layer of our mechanism comprises the web client, which serves as the interface for users to access content and issue commands to the system. By interacting with both users and the edge server, the web client facilitates communication and data transfer within the system. The web client also provides UI for control, management, data visualization, and service operations in the system, from deriving the optimal global model through FL to predicting worker safety in the outdoor construction environment.

Changes in data distribution among cross-silo edge nodes trigger the FL process. When there is a significant shift in data patterns or context within the monitored environments, such as a worker transitioning between different edge node environments, these modifications typically manifest. In such scenarios, the global model is periodically updated through an FL process. When nodes located in distinct silos observe and respond to changes in data distribution, they notify it to a central coordinating entity. When a significant change is identified, it serves as the impetus for initiating FL updates. The primary purpose of these FL updates is to accommodate the evolving data distribution by modifying the global model. In response to the trigger, edge nodes collaborate to refine the global model by incorporating their respective localized insights. This collaborative model training process guarantees the continued representativeness of the global model by assimilating the collective knowledge of all edge nodes, despite the possibility of time-varying data distribution fluctuations. For instance, data belonging to worker 1 is synchronized between Edge Nodes 1 and 2 as a result of their relocation in response to changes in data distribution. This synchronization mechanism ensures that local models are always up-to-date and that safety is continuously monitored. This method addresses the dynamic nature of monitored environments and ensures the continued relevance and effectiveness of the global model by periodically initiating FL updates in response to changes in data distribution. Therefore, the system retains its capacity to provide accurate and context-aware edge inference.

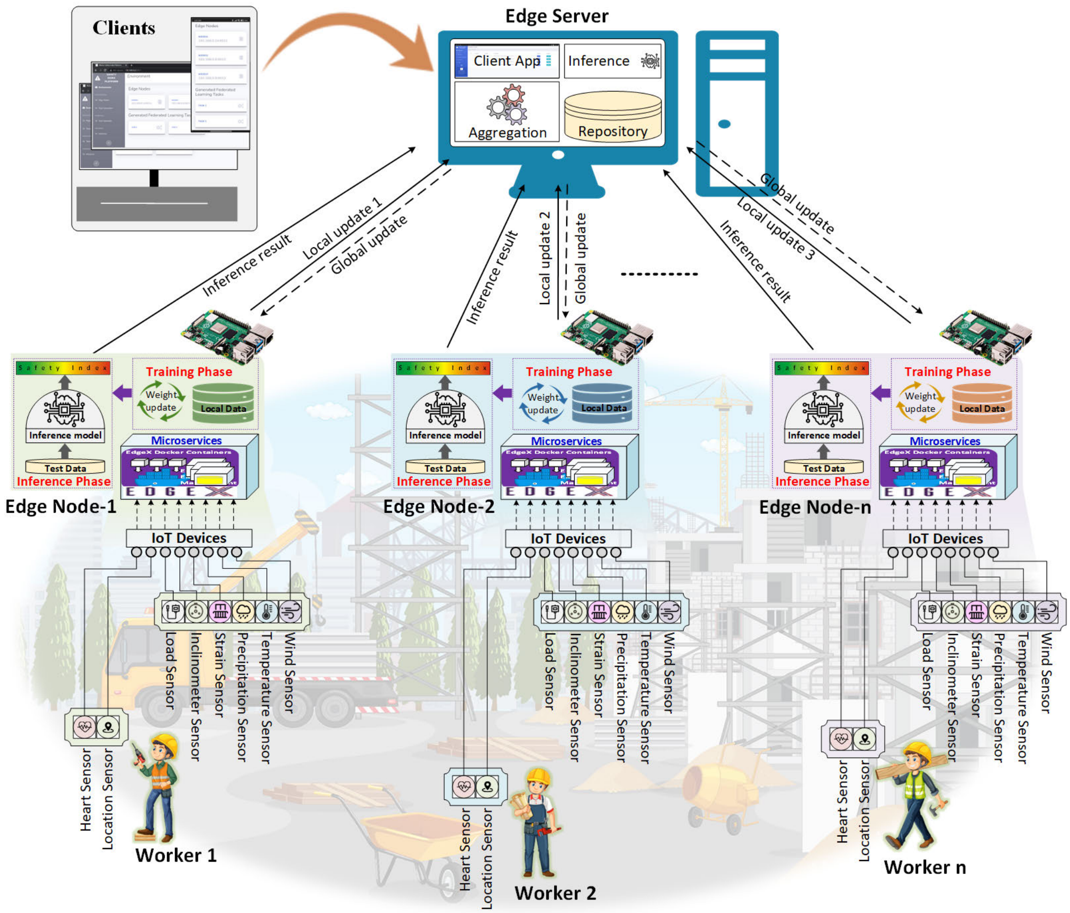


FIGURE 2. Proposed architecture for worker safety prediction mechanism.

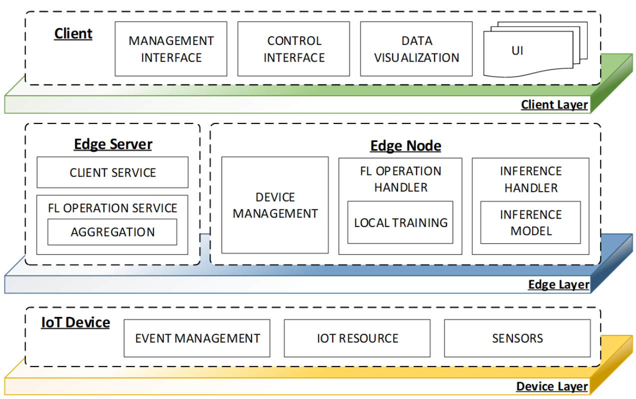


FIGURE 3. Layered architecture of the proposed worker safety prediction.

A. COLLABORATIVE TRAINING OF WORKER SAFETY PREDICTION MECHANISM

FL is a machine learning technique that allows for algorithm training utilizing local data samples dispersed over several decentralized edge devices or servers without the need to interact with the actual data. In our system, we utilize this technique to obtain the most optimal global model that efficiently predicts worker safety in the outdoor construction environment with high performance. During the FL process, the steps in each training round are as follows:

1) To begin the training process, the server initializes the global weights as the starting point during the first round. For subsequent rounds, the average weights of all nodes are used. The edge server sends global model weights to use for local training to the registered edge nodes it has chosen for the current round.
2) Participant nodes after receiving the global model adopt it and use their local data to train it. To facilitate training with DNN, the local data undergoes preprocessing and transformation to become a suitable format.
3) Nodes that took part in local training send the server their updated model parameters. Since each node uses diverse local data to train its model, the updated parameters will vary between nodes.
4) The server builds an improved global model by aggregating the parameters of the local model sent by each node. FedAVG is considered the main aggregation

approach. Up to convergence, the process is repeated from step 1.

5) Once the model has converged, the trained global model is distributed by the server to all participating edge nodes.
6) The optimal global model is deployed on all participating edge nodes to perform worker safety prediction. When using the global model for prediction, it is significant to remember that as in Step 2 of the model training process, the same preprocessing technique is still used to prepare the local data.

B. MODEL CONSOLIDATION FOR PROPOSED WORKER SAFETY PREDICTION MECHANISM

Typically, we can express the objective of FL as an optimization problem as follows:

$$\min l(w) = \sum_{j=1}^{J} \frac{n_j}{n} L_j(w)$$

$$\text{where } L_j(w) = \frac{1}{n_j} \sum_{i \in D_j}^{J} l_i(w) \tag{1}$$

Minimizing the loss, $l(w)$, over the data that are distributed, D, is the goal of the optimization problem, where the loss function of the global model is $l(w)$, the loss of j^{th} device is $L_j(w)$ and the loss for sample, i, is $l_i(w)$. The distribution of device's data points, j, is denoted by $D_j, j \in \{1, \ldots, J\}$, where $n_j = |D_j|$ is the size of D_j, and $n = \Sigma_{j=1}^{J} n_j$ is the total volume of data across all edge nodes. FedAVG is considered the main method of solving this optimization problem. FedAVG solves the optimization problem defined in FL by allowing devices to update their local models with multiple weight updates $w \longleftarrow w - \eta \nabla l(w)$ prior to submitting the server with the updated model weights. The details of model consolidation based on FedAVG algorithm for the proposed worker safety prediction mechanism [24] is presented in Algorithm 1.

Algorithm 1 Federated Averaging (FedAVG)

1: **EDGE_SERVER_UPDATE** (J, B, E, η) :
2: Initialize w_0
3: **for each round** $t = 1, 2, \ldots$ do
4: $m \leftarrow \max(C.J, 1)$
5: $S_t \leftarrow$ (random set of m edge nodes)
6: Broadcast global model weight to S_t edge nodes
7: **for** each edge node $j \in S_t$ **in parallel do**
8: $w_{t+1}^{j} \leftarrow$ *EDGE_NODE_UPDATE* (j, w_t)
9: Aggregate $w_{t+1} \leftarrow \sum_{j=1}^{J} \frac{n_j}{n} w_{t+1}^{j}$
10: **EDGE_NODE_UPDATE** (j, w) : ▷ Run on client j
11: $B \leftarrow$ (split D_j into batches B of size s_b)
12: **for** each local epoch i **from** 1 to E **do**
13: **for** batch $b \in B$ **do**
14: $w \longleftarrow w - \eta \nabla l(w)$
15: return w to server

Based on the fraction of edge nodes to perform computation on each round C, a random subset of edge nodes S_t is selected, and the global model weight is broadcasted to them at each round of FedAVG (lines 5 and 6). The selected edge nodes (line 8) then train their local models in parallel using their local data throughout several epochs. The **EDGE_NODE_UPDATE** procedure (line 10) presents the update process of the local model. D_j, denoted the local data is separated into batches B of size s_b (line 11), and the received global model is trained by the edge nodes with the created batches for multiple epochs (lines 12, 13, and 14). After each device in S_t computes its new local weights (line 15), these weights are transmitted to the edge server. The global model is updated by the edge server by computing the weighted average of the local weights received from the edge nodes (line 9). This process is repeated until the model converges (line 3), and the optimal model is finally deployed to all participants.

FedAVG is an alternative to Federated Stochastic Gradient Descent (FedSGD) that has shown significant improvements in time efficiency and communication [43]. Its fundamental concept is that averaging the weights is the same as averaging the gradients if all client nodes start from the same initialization parameters. As a result, the performance of the averaged model is not necessarily impacted. Furthermore, Adam optimizer is utilized in FedAVG in order to decrease the number of communication rounds and latency. The optimizer utilized in this context employs adaptive learning rates to hasten convergence and enhance learning efficacy when dealing with non-independent and identically distributed (non-IID) data [44].

C. OPERATIONAL OVERVIEW OF DEVELOPED WORKER SAFETY PREDICTION SYSTEM

Figure 4 shows the inclusive functional block diagram of the proposed mechanism. The system comprises four main functional blocks namely client, edge server, edge node, and IoT device. In the client block, UI is developed to enable users to control, manage, and operate the system via a web service. This interface facilitates interaction between the user and the system, ensuring a smooth user experience. In the edge server block, the client application service enables four major functions: node virtualization, FL task generation, FL operation, and inference operation. For each functional process, metadata is stored in the repository to ensure the efficient and effective operation of the system.

In the edge node block, four components are included:

- EdgeX core services, a framework that enables devices to securely and efficiently collect and transmit data.
- Device service for interconnecting IoT devices.
- FL operation handler for managing FL processes with the edge server.
- Inference handler for processing inference requests from the edge server.

The device block includes three blocks representing weather devices, worker devices, and building devices that are

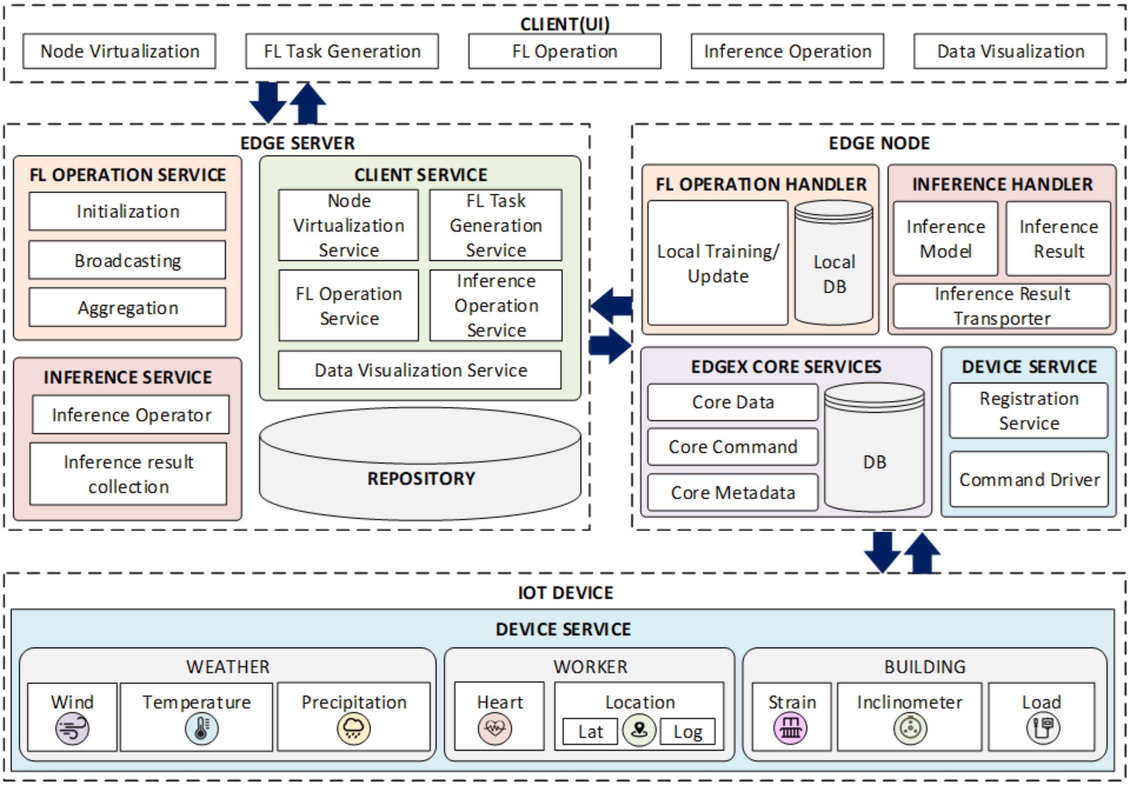


FIGURE 4. Functional block diagram of the proposed worker safety prediction approach.

registered in the EdgeX on the node through device service.

Figure 5 illustrates the general sequence diagram of the proposed worker safety prediction approach, which consists of several phases including system initialization, FL operation, and inference for predicting worker safety. In the system initialization phase, the user (web client) registers the edge node by providing its information such as name, IP address, and port number. After the node registration is complete, the user can generate a task for an FL operation by providing attributes such as name, communication round, model, and dataset. Once the system initialization phase is complete, the user can start the previously generated FL operation. Upon completion of the communication round and deployment of the final global model to the edge nodes, the FL phase concludes. In the inference phase, when the user requests to operate the inference, the edge server communicates with the edge node for inference. The edge node then retrieves data from IoT devices and performs data pre-processing and inference. The results of the inference are then transferred back to the server.

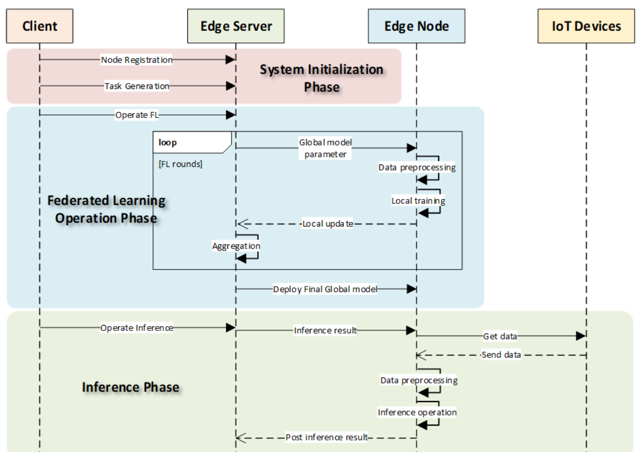


FIGURE 5. General sequence diagram of the proposed worker safety prediction approach.

IV. IMPLEMENTATION AND TESTING ENVIRONMENT RESULTS

A. IMPLEMENTATION SETUP

The following Figure 6 depicts our experimental environment architecture. For the edge server environment, the Ubuntu 20.04 desktop operating system is installed on a PC featuring an Intel® Core™ i9-11900F 2.5GHz CPU, 8 GB of memory, and 100 GB of storage, configured as a virtual machine. Each of the edge node environments is equipped with the Ubuntu 20.04 server operating system, installed on a Raspberry Pi 4B that features a quad-core 1.5GHz 64-bit CPU, 4GB of memory, and a 32GB MicroSD card. Each of the IoT device application environments is set up with the Ubuntu 20.04 desktop operating system on a PC that features an

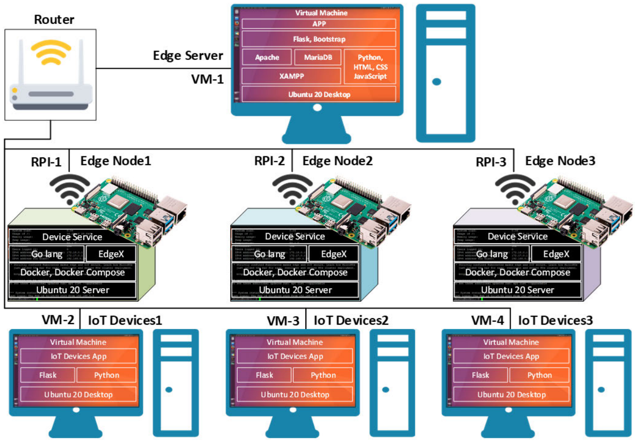


FIGURE 6. Experimental test-bed of federated learning-based worker safety prediction mechanism.

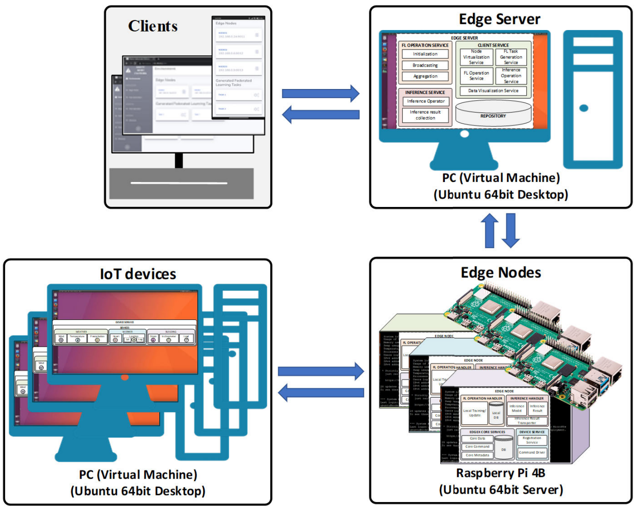


FIGURE 7. Experimental entities for worker safety prediction.

Intel® Core™ i9-11900F 2.5GHz CPU, 8 GB of memory, and 100 GB of storage, configured as a virtual machine. In our experimental environment, all devices are connected to a private HTTP network through a router. Virtual machines configured on the PC are connected via an Ethernet cable, which is bridged through an adapter in the Ubuntu operating system. Meanwhile, Raspberry Pi devices are connected via Wi-Fi.

The experimental entities of the proposed worker safety prediction mechanism are implemented on the respective platform to deliver the intended functionalities as depicted in the following Figure 7. Clients refer to the web-based interface that allows users to input and receive information, which is then delivered to the edge server. In our experiment, to allow access to the edge server via the internet, we utilize the Chrome Web Browser as the edge client on the PC. Client service application from the edge server provides UI. The contents of the UI are implemented using HTML, Bootstrap, and jQuery libraries with the Flask web framework.

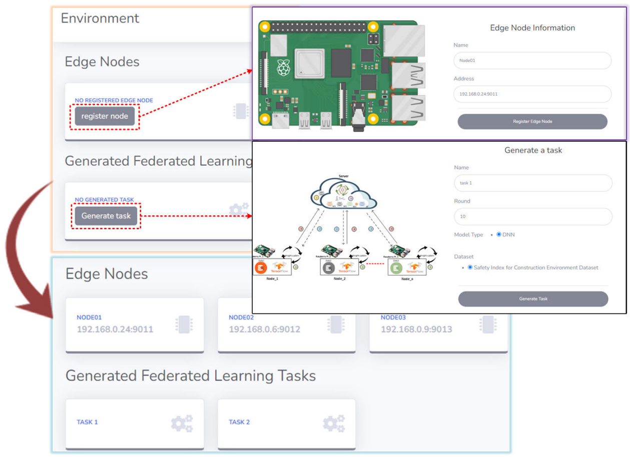


FIGURE 8. Initialization phase of the proposed worker safety prediction on UI.

Table 1 presents a detailed overview of the experimental environments for the IoT device, edge node, and edge server, including information on programming language, application (IDE), frameworks, libraries, and database.

B. INITIALIZATION PHASE OF THE PROPOSED WORKER SAFETY PREDICTION MECHANISM

Figure 8 depicts the initialization phase of the worker safety prediction mechanism based on FL on the Web Interface. In the domain of human-computer interaction in industrial design, UI refers to the area where interactions between humans and machines take place. Before the operation phase, the system needs to be initialized by registering the required number of client nodes to participate in training and generating desirable tasks to operate the FL. Hence, it can be seen that UI provides these two kinds of registration and it is showing that currently there are no registered nodes and no generated task with a sky blue highlighted square inside Figure 8. The purple highlighted rectangle inside Figure 8 shows the registration interface of an edge node (client node) on the edge server by providing its information. The desired number of the edge node can be registered by the unique name and its address (IP address and port number). In this image, it can be observed that edge node 1 (name with Node01) is being registered with its IP address information (192.168.0.24:9011). In our system, three edge nodes are used to conduct the experiment as we implemented three Raspberry Pi 4Bs as client nodes to be involved in FL. So, we register our three nodes by providing their names and addresses.

The black highlighted rectangle inside Figure 8 depicts the user interface for generating FL tasks during the initialization phase of worker safety prediction in the outdoor construction environment. To generate a task for the FL operation, the user can configure the attributes of the task such as task name, number of federated rounds, the type of model, and dataset as shown in the image. The sky blue highlighted rectangle

TABLE 1. Experimental environment of the proposed worker safety prediction approach.

Entity	Hardware	Software
Edge Server	PC, Ubuntu 20.04 Desktop (VM)	HTML 5, CSS3, Bootstrap, JavaScript, Python 3.8, PyCharm, TensorFlow 2.6, Flask 2.2.2, MySQL
Edge Node	Raspberry Pi 4 Model 4B, Ubuntu 20.04 Server	Python 3.8, Golang 3.6, Visual Studio Code, TensorFlow 2.6, Flask 2.2.2, EdgeX (Ireland), Docker 20.10.21, Docker Compose 1.25.0
IoT Device	PC, Ubuntu 20.04 Desktop (VM)	Python 3.8, Visual Studio Code, Flask 2.2.2

shows the result of the initialization phase of the proposed mechanism. The result depicts the information displayed on the user interface of three nodes registered and two FL tasks generated on the system.

C. OPERATION PHASE OF THE PROPOSED WORKER SAFETY PREDICTION MECHANISM

Figure 9 presents the operation phase of the proposed worker safety prediction mechanism. Users can operate one task at a time among the tasks generated on the system. It also shows the state of task 1 operating the FL with its attributes on the user interface. Task 1 comprises the attributes with 10 communication rounds of FL using the DNN model and local dataset name with worker safety for the outdoor construction environment dataset. To initiate the FL operation, users simply need to click the ''RUN'' button, which will update the button's color and text to indicate the operation has started on the system. Once the operation is complete, the button will return to its original state, and the user can review the FL task's history by clicking the ''History'' button.

The red dotted rectangle highlighted within Figure 9 depicts the history of the FL task displayed on UI. This task history encompasses metrics such as Mean Absolute Error (MAE), LOSS, Mean Absolute Percentage Error (MAPE), and training time for each node, and is organized by task name.

D. INFERENCE PHASE OF THE PROPOSED WORKER SAFETY PREDICTION MECHANISM

The following Figure 10 shows the inference environment on the user interface during the inference phase of the proposed worker safety prediction mechanism. Once the operation phase is done and after the trained global model is automatically distributed to the participant edge nodes, users can initiate the inference process. To begin the inference process, the user can easily click the ''Operate'' button. Once the button is clicked, the system will indicate that the operation has started by updating the button's text and color. After the completion of the inference operation, the button returns to its initial state. The user can click on the ''History'' button to view the history of the inference task. The highlighted red dotted rectangle shown in Figure 10 represents the history of the three times inference operation using the trained model by task 1 as displayed on UI. This task history records the predicted worker safety index and training time for each node and is arranged by task name.

V. PERFORMANCE ANALYSIS

A. DATASET AND EVALUATION METRICS

In this research study, safety data of construction workers is provided by a Research Institute, in the Republic of Korea. Table 2 presents a summary of the sensing devices used for data collection with their output descriptions. The dataset consists of 30,000 instances with 8 features. The study focuses on the outdoor construction environment and considers the following characteristics. (1)Weather conditions; The climate in the construction area plays a vital role. Extreme temperatures, precipitation, and wind speed are factors that can have an impact on construction materials, equipment, and the safety of workers [35], [45]. (2)Building conditions; In an outdoor construction environment, it is important to consider load, strain, [46] and inclinometer [47] measurements to ensure the structural integrity and stability of the building. (3)Worker conditions; Ensuring the well-being of workers and preventing accidents or health issues is paramount. Monitoring the heart rate of workers can provide insights into their physiological response to the work environment [48]. Additionally, their location information, particularly proximity to the danger zone, is crucial for ensuring their safety [49].

In order to devise a method for predicting worker safety, the data is simulated for three workers, each representing an edge node, and is validated on a server in an outdoor construction setting. Each row contains temperature, wind, precipitation, strain, inclinometer, load, heart rate, location, and safety information. Since worker safety is inferred from the knowledge of weather, building, and worker status, weather data is categorized into temperature, wind, and precipitation, building data into strain, inclinometer, and load, and worker status data into heart rate and location coordinates, encompassing latitude and longitude values.

The evaluation is performed on three edge nodes, two of which contained 9,000 rows of data and one containing 7,000 rows of data. Validation data consisting of 5,000 rows is used on the edge server. The significant diversity among the edge nodes in terms of features posed a challenge for capturing all the patterns among them with costly centralized training. To address this, we explored the use of FL to develop a unified global model for all edge nodes.

In this paper, we employed several statistical evaluation metrics to assess the effectiveness of our model, including MAE, Root Mean Square Error (RMSE), and Coefficient of Determination (R2) [50].

FIGURE 9. Operation phase of the proposed worker safety prediction on UI.

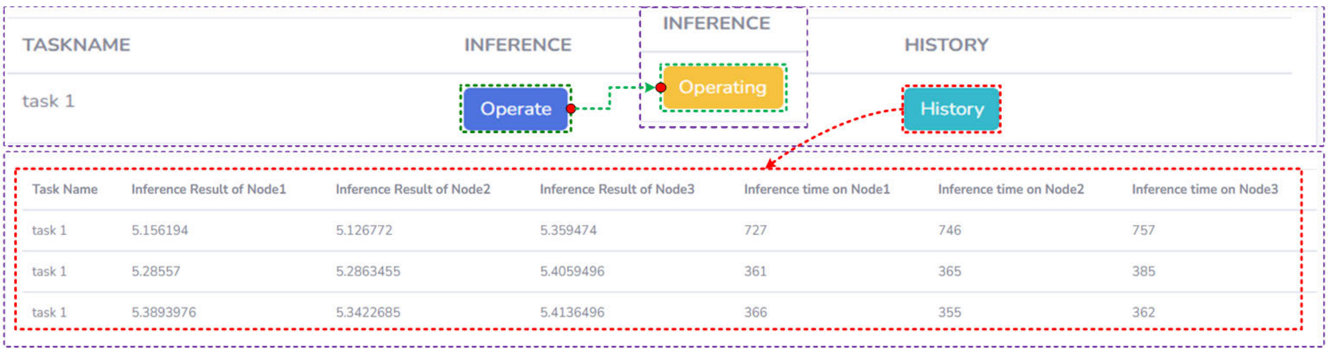


FIGURE 10. Inference phase of the proposed worker safety prediction on UI.

TABLE 2. Detailed descriptions of experimental sensors utilized for the proposed mechanism.

Parameter Name	Measuring Device	Output Description
Wind	Anemometer	Wind velocity (meter per second)
Precipitation	Pluviometer	Height of rainfall (millimeters)
Temperature	DHT22	Outdoor temperature (Degrees Celcius)
Load	Load Cell	Pulling force (Kilo Newton)
Strain	Strain gauge	Change in length (Microstrain)
Inclinometer	Inclinometer	Slope of object (Degree)
Heart Rate	Heart Rate Sensor	Human heartbeat (Beats per minute)
Location	Handheld GPS	Proximity to the danger zone (Meter)

- MAE: A commonly used evaluation metric that quantifies the deviation between the actual and predicted values of a variable. The range of MAE varies from 0 to ∞. A value of 0 indicates the best possible performance, with predictions perfectly matching actual values. Increasing MAE signifies higher discrepancies between the model's predictions and the actual data. It is calculated as the absolute difference between the predicted and target values. It is given by equation 2 [51]:

$$MAE = \frac{\sum_{i=1}^{n} \left| Y_i - \hat{Y}_i \right|}{n} \tag{2}$$

- RMSE: The objective of employing this metric is to assess the regression model's error rate and confirm that the error magnitudes are comparable to those of the targets. RMSE range spans from 0 to ∞. A value of '0'

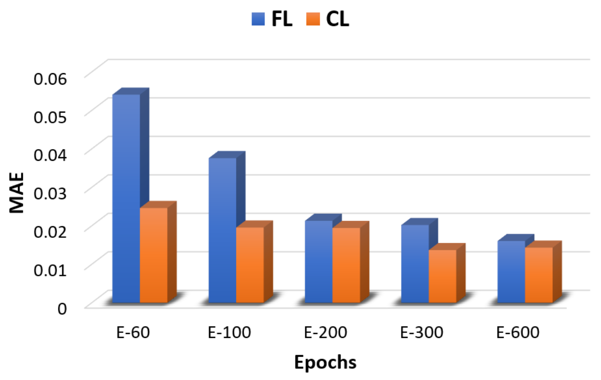


FIGURE 11. MAE comparison between FL and CL.

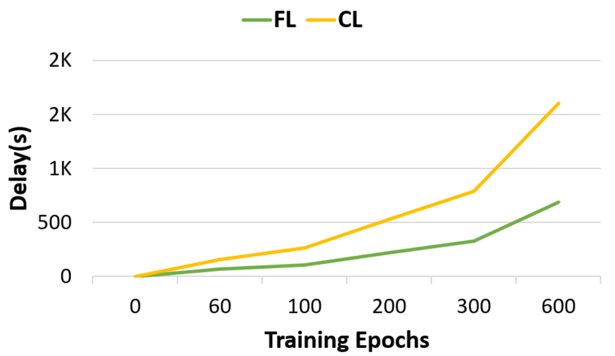


FIGURE 12. Training delay comparison between FL and CL.

indicates the regression model's accurate performance on unseen data, while a higher RMSE value suggests a significant error in the prediction process. It is computed as the square root of the Mean Square Error (MSE), which is defined by the following equation 3 [51]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)\Big|^2}{n}} \tag{3}$$

- Coefficient of Determination (R2): A statistical metric that indicates the percentage of variability in the dependent variable that can be explained by one or more independent variables. The R2 score ranges from 0 to 1, where a value of 0 indicates the regression model's weakest performance and a value of 1 signifies its optimal performance over unseen data samples. The equation for R2 is 4 [52]:

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}} \tag{4}$$

Although our primary focus is on distributed machine learning at the edge, we also conduct analysis regarding the error, training time, and inference time of our work compared to centralized machine learning. In Figure 11, we present a comparison of error, measured by the MAE, between FL and centralized learning (CL). The findings indicate that CL is better than FL in every training epoch. However, the error gap between FL and CL diminishes at 600 training epochs, suggesting the potential for improving FL's performance by increasing the number of training epochs. Besides, FL stands out for its strong emphasis on security, ensuring data privacy is safeguarded by eliminating the need to transmit local data to a central server during the training process.

Additionally, in Figure 12, we present a comparison of the training time between FL and CL, measured in seconds (s). This experiment is conducted on the constrained device, Raspberry Pi 4B. Overall, FL demonstrates significantly lower training time compared to CL, owing to its distributed training approach.

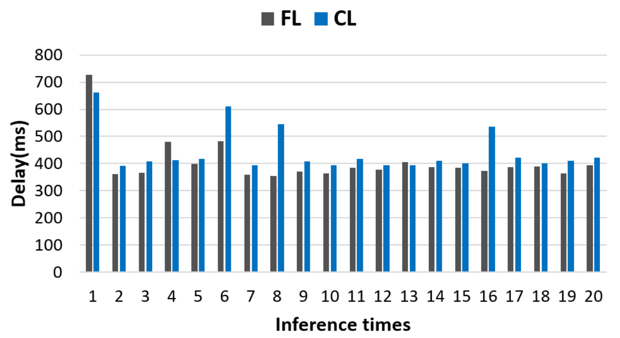


FIGURE 13. Inference delay comparison between FL and CL.

TABLE 3. Statistical Performance Evaluation of the worker safety prediction model.

FL Round	MAE	RMSE	R2
FL-3	0.053	0.077	0.712
FL-5	0.038	0.062	0.816
FL-10	0.028	0.046	0.896
FL-15	0.024	0.043	0.909
FL-30	**0.015**	**0.025**	**0.969**

Moreover, Figure 13 presents a comparison of the inference time between FL and CL, measured in milliseconds (ms). In general, FL demonstrates shorter inference times as the model is deployed at the edge following the communication rounds, resulting in faster inference performance.

Furthermore, Table 3 provides a summary of the statistical performance evaluation for the predicted model, based on the number of federated rounds. All evaluation results were obtained from the DNN learning model trained with 20 epochs on the clients. The table reveals that the learning model's performance improves as the federated round and training epoch number increase, as demonstrated by the evaluation metrics of 0.015 MAE, 0.025 RMSE, and 0.969 R2.

Figure 14 is used to illustrate the evaluation of the DNN learning model using the MAE metric. By comparing the

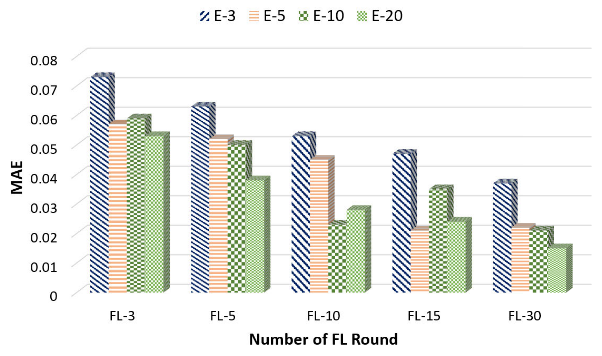


FIGURE 14. MAE evaluation of the learning model.

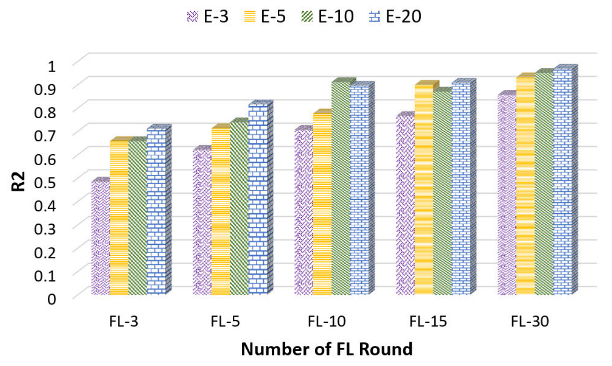


FIGURE 16. R2 evaluation of the learning model.

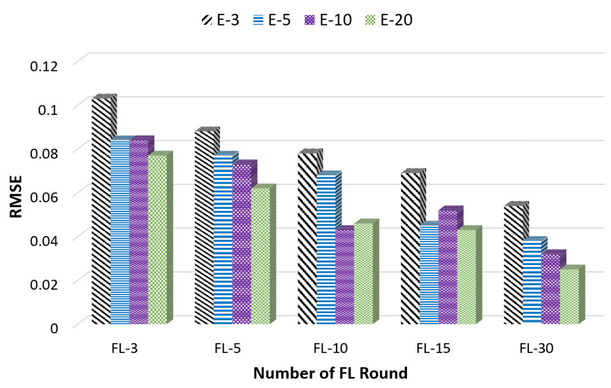


FIGURE 15. RMSE evaluation of the learning model.

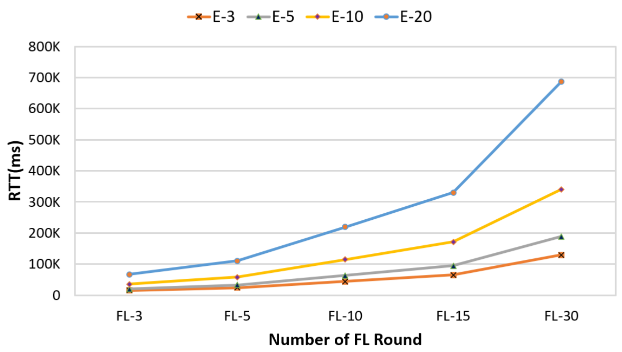


FIGURE 17. RTT based on the federated round and training epoch number.

results obtained from different numbers of FL rounds and training epochs on the node, it is apparent from the results that as the federated round and training epoch number increase, the error rate or loss value decreases. The optimal performance was attained using 30 federated rounds and 20 training epochs.

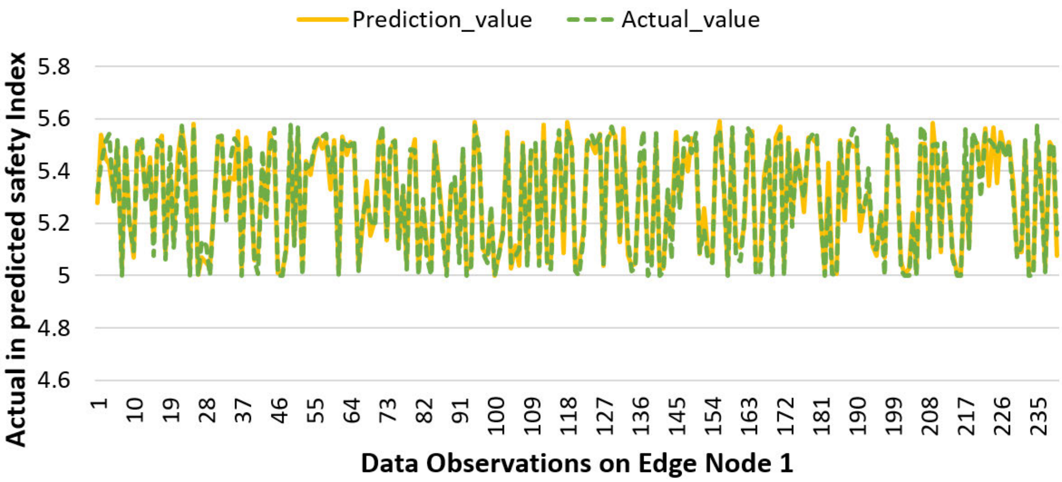
Additionally, the evaluation of the DNN learning model based on the RMSE metric is presented in Figure 15. The results obtained from various numbers of FL rounds and training epochs on the edge node reveal a decreasing error rate (loss value) as the federated round and training epoch number increase. Optimal performance was achieved with 30 federated rounds and 20 training epochs.

Also, the following Figure 16 depicts the evaluation of the DNN learning model using the R2 metric. Analysis of the results obtained from different numbers of FL rounds and training epochs on the node indicates a reduction in the error rate (loss value) as the number of federated rounds and training epochs increases. The best performance is observed with 30 federated rounds and 20 training epochs.

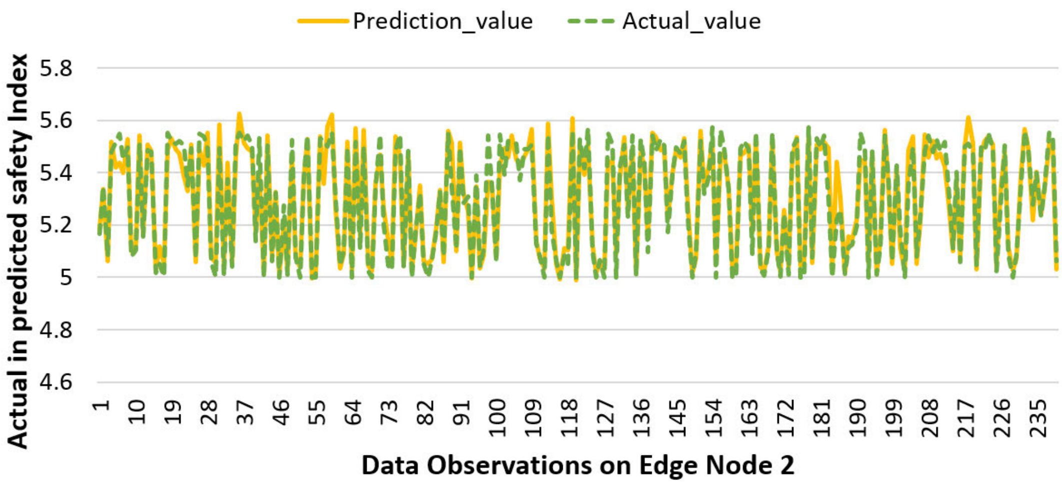
In addition, the round-trip time (RTT) measured in milliseconds (ms) is shown in Figure 17, which serves as an indicator of the latency in each communication round, with respect to the number of federated rounds and training epochs. The figure illustrates that RTT experiences an increase as the number of federated rounds and the epochs of the training increase. However, during the first three rounds of FL, all training epochs had a lower latency with values of less than 100k ms.

Also, the succeeding Figure 18(a)(b)(c) illustrates the relationship between the actual and predicted worker safety index predicted on each node using the optimal predicted model obtained from the DNN learning model, which was trained with 30 federated rounds and 20 training epochs. For simulating one-hour-ahead predictions of the safety of a worker in an outdoor construction environment, the data points for prediction were collected every 15 seconds, with 240 data observations, which is reasonable in fatal environments like construction sites, although shorter time intervals for prediction should be considered to improve the worker's safety. From the segment depicted in the figure, it is noticeable that the predicted values align closely with the actual values.

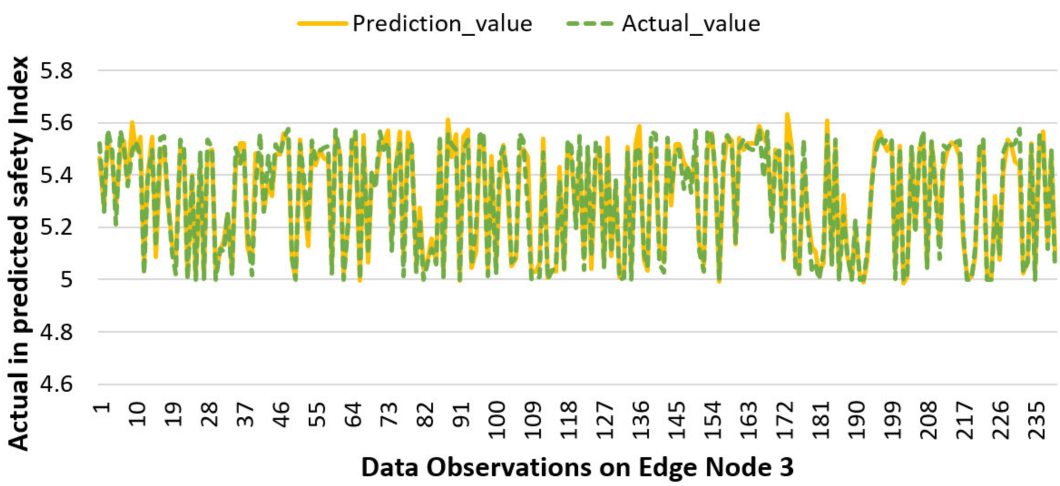
Moreover, Figure 19 displays the time delay, measured in milliseconds (ms), in predicting worker safety on three nodes. The figure compares the results of 20 prediction delays collected from each node. Initially, an average delay of approximately 700 ms is observed for inference on each node, whereas the subsequent prediction times took around 400 ms on average.



(a) Actual versus predicted worker safety index for edge node 1



(b) Actual versus predicted worker safety index for edge node 2



(c) Actual versus predicted worker safety index for edge node 3

FIGURE 18. Comparison of actual and predicted worker safety index for one-hour ahead prediction on three edge nodes.

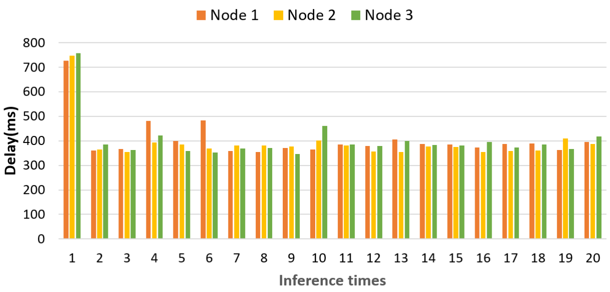


FIGURE 19. Time delay of predicting worker safety on three nodes.

VI. CONCLUSION

Construction industry has been remarked as one of the most hazardous industries accounting for more work-related injuries and deaths than any other industry in the world. Therefore, an effective construction safety management system is required to reduce damage caused by construction site accidents. To create a safe construction environment by reducing potential damage and hazardous situations, we proposed the worker safety prediction mechanism based on the FL technique-assisted edge intelligence in the outdoor construction environment. To achieve an efficient and adaptive predictive model, this paper integrates FL, which offers benefits such as preserving data privacy and leveraging rich resources by aggregating the model's parameters through training local data distributed among the edge nodes. To enhance the use of IoT devices in the construction environment, we also utilized the EdgeX framework that offers capabilities such as conveniently managing devices and data and interoperability between devices and applications at the IoT Edge. Additionally, we provide a web service that interacts with the system through UI to easily manage, operate, and assess for the sake of worker safety in the outdoor construction environment. As far as we know, no active research is still done for predicting worker safety in the outdoor construction environment while integrating with the FL technique to achieve the efficient and effective predicting model in Edge. It is expected that the proposed worker safety prediction mechanism can effectively seek worker safety in the outdoor construction environment while unleashing the capacity of cutting-edge technologies.

In the future, we intend to expand our research to experiment with real-world data, while considering additional features to ensure worker safety in outdoor construction environments. Furthermore, we will conduct comparative research with other models to enhance performance.

REFERENCES

[1] *68% of the World Population Projected to Live in Urban Areas by 2050, Says UN*. Accessed: May 4, 2023. [Online]. Available: https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html

[2] M. Zhang, R. Shi, and Z. Yang, "A critical review of vision-based occupational health and safety monitoring of construction site workers," *Saf. Sci.*, vol. 126, Jun. 2020, Art. no. 104658.

[3] *Revealed: 6,500 Migrant Workers Have Died in Qatar Since World Cup Awarded*. Accessed: May 4, 2023. [Online]. Available: https://www.theguardian.com/global-development/2021/feb/23/revealed-migrant-worker-deaths-qatar-fifa-world-cup-2022

[4] N. Jayanthi, "IoT based-civil labour safety monitoring system in construction site," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 9, pp. 1723–1728, 2021.

[5] S. H. Kim, H. G. Ryu, and C. S. Kang, "Development of an IoT-based construction site safety management system," in *Proc. Int. Conf. Inf. Sci. Appl.* Cham, Switzerland: Springer, 2019, pp. 617–624.

[6] N. Petrović and Đ. Kocić, "IoT-based system for COVID-19 indoor safety monitoring," in *Proc. IcETRAN*, 2020.

[7] N. Singh, V. K. Gunjan, G. Chaudhary, R. Kaluri, N. Victor, and K. Lakshmanna, "IoT enabled HELMET to safeguard the health of mine workers," *Comput. Commun.*, vol. 193, pp. 1–9, Sep. 2022.

[8] N. Mishra and S. Silakari, "Predictive analytics: A survey, trends, applications, oppurtunities & challenges," *Int. J. Comput. Sci. Inf. Technol.*, vol. 3, no. 3, pp. 4434–4438, 2012.

[9] T. Falatouri, F. Darbanian, P. Brandtner, and C. Udokwu, "Predictive analytics for demand forecasting—A comparison of SARIMA and LSTM in retail SCM," *Proc. Comput. Sci.*, vol. 200, pp. 993–1003, Jan. 2022.

[10] A. A. Koelmans, P. E. Redondo-Hasselerharm, N. H. M. Nor, V. N. de Ruijter, S. M. Mintenig, and M. Kooi, "Risk assessment of microplastic particles," *Nature Rev. Mater.*, vol. 7, no. 2, pp. 138–152, Jan. 2022.

[11] G. Zhang, Z. Li, J. Huang, J. Wu, C. Zhou, J. Yang, and J. Gao, "EFraudCom: An e-commerce fraud detection system via competitive graph neural networks," *ACM Trans. Inf. Syst.*, vol. 40, no. 3, pp. 1–29, Jul. 2022.

[12] E. W. T. Ngai and Y. Wu, "Machine learning in marketing: A literature review, conceptual framework, and research agenda," *J. Bus. Res.*, vol. 145, pp. 35–48, Jun. 2022.

[13] X. Su, X. Yan, and C.-L. Tsai, "Linear regression," *Wiley Interdiscipl. Rev., Comput. Statist.*, vol. 4, no. 3, pp. 275–294, 2012.

[14] M. P. LaValley, "Logistic regression," *Circulation*, vol. 117, no. 18, pp. 2395–2399, 2008.

[15] L. Rokach and O. Maimon, "Decision trees," in *Data Mining and Knowledge Discovery Handbook*. Boston, MA, USA: Springer, 2005, pp. 165–192.

[16] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Jul. 1998.

[17] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[18] R. Yi Man Li, L. Song, B. Li, M. James C. Crabbe, and X.-G. Yue, "Predicting carpark prices indices in Hong Kong using AutoML," *Comput. Model. Eng. Sci.*, vol. 134, no. 3, pp. 2247–2282, 2023.

[19] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[20] N. Prasad, B. Rajpal, K. K. R. Mangalore, R. Shastri, and N. Pradeep, "Frontal and non-frontal face detection using deep neural networks (DNN)," *Int. J. Res. Ind. Eng.*, vol. 10, no. 1, pp. 9–21, 2021.

[21] E. Fathi and B. M. Shoja, "Deep neural networks for natural language processing," in *Handbook of Statistics*, vol. 38. Amsterdam, The Netherlands: Elsevier, 2018, pp. 229–316.

[22] G. Singh, M. Pal, Y. Yadav, and T. Singla, "Deep neural network-based predictive modeling of road accidents," *Neural Comput. Appl.*, vol. 32, no. 16, pp. 12417–12426, Aug. 2020.

[23] L. Das, A. Sivaram, and V. Venkatasubramanian, "Hidden representations in deep neural networks: Part 2. Regression problems," *Comput. Chem. Eng.*, vol. 139, Aug. 2020, Art. no. 106895.

[24] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial Intelligence and Statistics*. PMLR, 2017, pp. 1273–1282.

[25] F. Luo, R. Y. M. Li, M. J. C. Crabbe, and R. Pu, "Economic development and construction safety research: A bibliometrics approach," *Saf. Sci.*, vol. 145, Jan. 2022, Art. no. 105519.

[26] B. Sherafat, C. R. Ahn, R. Akhavian, A. H. Behzadan, M. Golparvar-Fard, H. Kim, Y.-C. Lee, A. Rashidi, and E. R. Azar, "Automated methods for activity recognition of construction workers and equipment: State-of-the-art review," *J. Construct. Eng. Manage.*, vol. 146, no. 6, Jun. 2020, Art. no. 03120002.

[27] R. Y. M. Li and R. Y. M. Li, "Three generations of construction safety informatics: A review," in *Construction Safety Informatics*. Singapore: Springer, 2019, pp. 1–12.

[28] J. Häikiö, J. Kallio, S.-M. Mäkelä, and J. Keränen, "IoT-based safety monitoring from the perspective of construction site workers," *Int. J. Occupational Environ. Saf.*, vol. 4, no. 1, pp. 1–14, Apr. 2020.

[29] R. Y. M. Li, B. Tang, and K. W. Chau, "Sustainable construction safety knowledge sharing: A partial least square-structural equation modeling and a feedforward neural network approach," *Sustainability*, vol. 11, no. 20, p. 5831, Oct. 2019.

[30] M. Khan, R. Khalid, S. Anjum, N. Khan, S. Cho, and C. Park, "Tag and IoT based safety hook monitoring for prevention of falls from height," *Autom. Construct.*, vol. 136, Apr. 2022, Art. no. 104153.

[31] R. Edirisinghe, "Digital skin of the construction site: Smart sensor technologies towards the future smart construction site," *Eng., Construct. Architectural Manag.*, vol. 26, no. 2, pp. 184–223, 2019.

[32] I. Awolusi, E. Marks, and M. Hallowell, "Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices," *Autom. Construct.*, vol. 85, pp. 96–106, Jan. 2018.

[33] J. Teizer, H. Neve, H. Li, S. Wandahl, J. König, B. Ochner, M. König, and J. Lerche, "Construction resource efficiency improvement by long range wide area network tracking and monitoring," *Autom. Construct.*, vol. 116, Aug. 2020, Art. no. 103245.

[34] R. Kanan, O. Elhassan, and R. Bensalem, "An IoT-based autonomous system for workers' safety in construction sites with real-time alarming, monitoring, and positioning strategies," *Autom. Construct.*, vol. 88, pp. 73–86, Apr. 2018.

[35] K. Kang and H. Ryu, "Predicting types of occupational accidents at construction sites in Korea using random forest model," *Saf. Sci.*, vol. 120, pp. 226–236, Dec. 2019.

[36] M. Jahangiri, H. R. J. Solukloei, and M. Kamalinia, "A neuro-fuzzy risk prediction methodology for falling from scaffold," *Saf. Sci.*, vol. 117, pp. 88–99, Aug. 2019.

[37] J. Choi, B. Gu, S. Chin, and J.-S. Lee, "Machine learning predictive model based on national data for fatal accidents of construction workers," *Autom. Construct.*, vol. 110, Feb. 2020, Art. no. 102974.

[38] M. Wang, P. Wong, H. Luo, S. Kumar, V. Delhi, and J. Cheng, "Predicting safety hazards among construction workers and equipment using computer vision and deep learning techniques," in *Proc. Int. Symp. Automat. Robot. Construct.*, 2019, pp. 399–406.

[39] A. Vaid et al., "Federated learning of electronic health records to improve mortality prediction in hospitalized patients with COVID-19: Machine learning approach," *JMIR Med. Informat.*, vol. 9, no. 1, Jan. 2021, Art. no. e24207.

[40] T. Borger, P. Mosteiro, H. Kaya, E. Rijcken, A. A. Salah, F. Scheepers, and M. Spruit, "Federated learning for violence incident prediction in a simulated cross-institutional psychiatric setting," *Expert Syst. Appl.*, vol. 199, Aug. 2022, Art. no. 116720.

[41] X. Li, H.-L. Chi, W. Lu, F. Xue, J. Zeng, and C. Z. Li, "Federated transfer learning enabled smart work packaging for preserving personal image information of construction worker," *Autom. Construct.*, vol. 128, Aug. 2021, Art. no. 103738.

[42] T. Qi, L. Chen, G. Li, Y. Li, and C. Wang, "FedAGCN: A traffic flow prediction framework based on federated learning and asynchronous graph convolutional network," *Appl. Soft Comput.*, vol. 138, May 2023, Art. no. 110175.

[43] D. Chai, L. Wang, L. Yang, J. Zhang, K. Chen, and Q. Yang, "FedEval: A holistic evaluation framework for federated learning," 2020, *arXiv:2011.09655*.

[44] X. Li, K. Huang, W. Yang, S. Wang, and Z. Zhang, "On the convergence of FedAvg on non-IID data," 2019, *arXiv:1907.02189*.

[45] S. Karthick, S. Kermanshachi, A. Pamidimukkala, and M. Namian, "A review of construction workforce health challenges and strategies in extreme weather conditions," *Int. J. Occupational Saf. Ergonom.*, vol. 29, no. 2, pp. 773–784, Apr. 2023.

[46] S. Sakhakarmi, J. Park, and C. Cho, "Enhanced machine learning classification accuracy for scaffolding safety using increased features," *J. Construct. Eng. Manage.*, vol. 145, no. 2, Feb. 2019, Art. no. 04018133.

[47] T.-Y. Liu, S.-J. Ho, H.-P. Tserng, and H.-K. Tzou, "Using a unique retaining method for building foundation excavation: A case study on sustainable construction methods and circular economy," *Buildings*, vol. 12, no. 3, p. 298, Mar. 2022.

[48] S. Anwer, H. Li, M. F. Antwi-Afari, W. Umer, and A. Y. L. Wong, "Evaluation of physiological metrics as real-time measurement of physical fatigue in construction workers: State-of-the-art review," *J. Construct. Eng. Manage.*, vol. 147, no. 5, May 2021, Art. no. 03121001.

[49] S. Dong, H. Li, and Q. Yin, "Building information modeling in combination with real time location systems and sensors for safety performance enhancement," *Saf. Sci.*, vol. 102, pp. 226–237, Feb. 2018.

[50] E. W. Steyerberg, A. J. Vickers, N. R. Cook, T. Gerds, M. Gonen, N. Obuchowski, M. J. Pencina, and M. W. Kattan, "Assessing the performance of prediction models: A framework for traditional and novel measures," *Epidemiology*, vol. 21, no. 1, pp. 128–138, 2010.

[51] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not," *Geoscientific Model Develop.*, vol. 15, no. 14, pp. 5481–5487, Jul. 2022.

[52] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, p. e623, Jul. 2021.

SA JIM SOE MOE received the B.S. degree in computer technology from the University of Computer Studies, Yangon, in 2010. He is currently pursuing the M.S. degree in computer engineering with Jeju National University, South Korea. His research interests include machine learning, edge computing, and the development of the Internet of Things.

BONG WAN KIM received the B.S. degree in electronics engineering from Han-yang University, South Korea, in 1992, the M.S. degree in electrical engineering from Korea Advanced Institute of Science and Technology (KAIST) in 1994, and the Ph.D. degree in computer science and electrical engineering from KAIST in 2000. He is a Principal Researcher at Electronics and Telecommunications Research Institute (ETRI), working on development of edge computing system for solving urban problems. He is currently interested in edge computing, wireless sensor network, and data analysis based on artificial intelligence.

ANAM NAWAZ KHAN received the B.S. and M.S. degrees in computer science from COMSATS University Islamabad, Islamabad, Pakistan, in 2016 and 2019, respectively. She joined various reputable academic institutions and later decided to expand her knowledge by continuing her academic career. In 2020, she moved to the Republic of Korea and started working as a Ph.D. Research Fellow with the Department of Computer Engineering, Jeju National University, Jeju, South Korea. She is currently working on integrating the AI-enabled IoT platforms and edge computing to provide autonomous intelligent services for smart environments. Her research interests include development and optimization of artificial intelligence and machine learning-based applications in smart spaces.

XU RONGXU received the B.S. degree in computer science from the Yanbian University of Science and Technology, China, in 2014, the M.S. degree in computer engineering from Konkuk University, South Korea, in 2017, and the Ph.D. degree from the Mobile Computing Laboratory (MCL), Department of Computer Engineering, Jeju National University, South Korea, in 2022. He has been a Postdoctoral Researcher with the Big Data Research Center, Jeju National University, since March 2022. In addition, he studies optimization and machine learning for smart homes, DNN, and FL for edge intelligent practice. He is currently a Project Manager for a research project supported by the National Research Foundation of Korea. His research interest includes edge computing.

NGUYEN ANH TUAN received the B.S. degree in applied mathematics and informatics from the Hanoi University of Science and Technology, in 2021. He is currently pursuing the integrated master's and Ph.D. degree in computer engineering with Jeju National University, South Korea. His research interests include the development of the Internet of Things, optimization, and machine learning.

KWANGSOO KIM received the B.S. degree in information engineering and the M.S. degree in computer science from Korea University, South Korea, in 1993 and 1995, respectively, and the Ph.D degree in computer engineering from Chungnam National University, South Korea, in 2016. Since 1995, he has been worked as a Principal Researcher with City and Spatial ICT Research Team, Electronics and Telecommunications Research Institute (ETRI), South Korea. His research interests include spatial information, geographic information system, location based services, sensor networks, and IoT platforms.

DO HYEUN KIM received the B.S. degree in electronics engineering and the M.S. and Ph.D. degrees in information telecommunication from Kyungpook National University, South Korea, in 1988, 1990, and 2000, respectively.

He joined the Agency of Defense Development (ADD), from March 1990 to April 1995. From 2008 to 2009, he was a Visiting Researcher with the Queensland University of Technology, Australia. Since 2004, he has been with Jeju National University, South Korea, where he is currently a Professor with the Department of Computer Engineering. His research interests include sensor networks, M2M/IoT, energy optimization and prediction, intelligent service, and mobile computing.