



A Federated Interactive Learning IoT-Based Health Monitoring Platform

Sadi Alawadi¹(✉), Victor R. Kebande², Yuji Dong³, Joseph Bugeja⁴,
Jan A. Persson⁴, and Carl Magnus Olsson⁴

¹ Department of Information Technology, Uppsala University, Uppsala, Sweden
sadi.alawadi@it.uu.se

² Department of Computer Science, Electrical and Space Engineering,
Luleå University of Technology, Luleå, Sweden
victor.kebande@ltu.se

³ School of Internet of Things, Xi'an Jiaotong-Liverpool University, Suzhou, China
yuji.dong02@xjtlu.edu.cn

⁴ Department of Computer Science, Malmö University, Malmö, Sweden
{[joseph.bugeja](mailto:joseph.bugeja@mau.se), [jan.a.persson](mailto:jan.a.persson@mau.se), [carl.magnus.olsson](mailto:carl.magnus.olsson@mau.se)}@mau.se

Abstract. Remote health monitoring is a trend for better health management which necessitates the need for secure monitoring and privacy-preservation of patient data. Moreover, accurate and continuous monitoring of personal health status may require expert validation in an active learning strategy. As a result, this paper proposes a Federated Interactive Learning IoT-based Health Monitoring Platform (FIL-IoT-HMP) which incorporates multi-expert feedback as ‘Human-in-the-loop’ in an active learning strategy in order to improve the clients’ Machine Learning (ML) models. The authors have proposed an architecture and conducted an experiment as a proof of concept. Federated learning approach has been preferred in this context given that it strengthens privacy by allowing the global model to be trained while sensitive data is retained at the local edge nodes. Also, each model’s accuracy is improved while privacy and security of data has been upheld.

Keywords: IoT · Healthcare · Federated · Machine learning

1 Introduction

Continuous advancement of the Internet of Things (IoT) healthcare systems has been experienced as a result of the sporadic technological changes, particularly in IoT device proliferation, and the need to manage the ever-rising quantity of patient data. Notably, these proliferation have allowed the usage of several health devices like wearable sensors that are able to measure and monitor several personal health parameters, which in some situations are able to create a trigger mechanism in case of a potential health incident. In order to make an accurate

prognosis using the data from IoT devices, most related healthcare systems currently leverages machine learning approaches for purposes of making decisions automatically. While this has seen an improved diagnosis and efficient detection of diseases [1], there have been several limitations such as lack of annotated data used to train the ML models, which in this context makes IoT-health systems unreliable and ineffective. Also, ineffective and malevolent coordination of ML model may lead to potential attacks and data leakage. In particular cases, this may lead to privacy infringement of patient data, which on similar situations puts the security of data at risk, hence creating mistrust among different parties. Indeed, in a recent study, Ponemon Institute identified that health data is the most targeted by cybercriminals [2] and that attacks on IoT devices were reported to be increasing by three-fold in 2018 [3].

Therefore, to improve both the ML model performance and accuracy in order to make IoT-health systems reliable and effective, there is need to interactively incorporate the expert's domain knowledge as 'human-in-the loop' to help in providing heuristic-based knowledge of the system while the IoT health system learns, and the need to preserve privacy, security and trust. Hence, Federated Machine Learning (FML) [4,5] will preserve data privacy by training the ML model over the user data locally without moving the data. In this context, ML model still can be adapted or contextualized locally which is more effective as opposed to leveraging a single trained model. Moreover, all the edge nodes will participate in training the ML model collaboratively using their data. Based on that fact, all ML models will share their learned knowledge among all participant nodes. To bring out the problem that is being addressed in this paper, we consider the following scenario:

The number of elderly people is on the rise and quite a good number of the them prefer to live in their homes (houses/apartments) devoid of privacy violations. However, in some cases, older people with chronic diseases are susceptible to other diseases like heart attacks or accidental falls etc. Monitoring their health remotely without human intervention using an IoT-based devices offers a suitable solution in this case. That notwithstanding, as the sensing data is massive and continuously generated, it may be impossible for humans to continually and accurately monitor and explore this data. A suitable solution is to use ML approaches to classify the sensed data into different events and let the domain experts only to validate those data deemed to have important events, for example, de-identification etc. Also, this kind of approach faces formidable challenges and issues. For example, the ML model's quality is paramount when it is required to give accurate classifications. A wrong classification or misdiagnosis could lead to serious consequences. Simultaneously, a good ML model needs massive data and many medical experts, which is probably impossible or too costly for a company. Additionally, enforcing the privacy and security of the data should be a priority in this context.

The authors take a step in addressing the aforementioned scenario by applying an IoT-based health monitoring platform with federated learning and interactive learning strategy that allows the knowledge from ML models that is trig-

gered by the domain experts to be shared, at the same time the clients are able to reap the benefits of such domain knowledge given the accuracy of these ML models.

The remainder of this paper is structured as follows: Related work is presented in Sect. 2, while the system architecture is presented in Sect. 3. This is followed by experiments in Sect. 4 and a conclusion in Sect. 5.

2 Related Work

A thought that a human may instinctively outperform a machine learning algorithm has been explored based on existing evidence on the diagnostic radiologic image. This is represented as a suitable approach that solves the expert-in the loop technique [6]. However, while it looks relevant, its effectiveness is rarely investigated when it is mapped to the patients' privacy. Also, an architecture that acts as a remote human-in-the-loop named SENS-U allows health monitoring for Wireless Body Sensor Network (WBSN) for patients. It is able to monitor terminals of medical centers via four body vital signs for personal healthcare [7]. A cost-optimal multi-expert (Co-MEAL) approach that has machine learning adaptability allows the machine learning model to be able to learn from a variety of experts, e.g., the human oracle or a digital device. The advantage of this process is that it reduces the cost of labeling data while it capitalizes on the collaboration among experts with the main aim of enriching the knowledge [8]. Based on the expert selection module of the Co-MEAL, a collaborative-multi-expert architecture by [9] has been designed to be able to manage knowledge from heterogeneous sources by incorporating a technique that allows experts to collaborate in order to increase their knowledge. Additionally, this work, was able to propose an expert selection algorithm that could be applied in a real world scenario by utilising active and transfer learning strategy, while the expert selection is executed as an expert unit in the Co-MEAL architecture. Also, a FedHealth framework that utilises transfer learning has been able to build personalized models through activity recognition experiments. Based on this study, accurate healthcare is achieved by FedHealth while at the same time privacy and security is upheld. When federated learning is used an accuracy of 99.4% is achieved as opposed to 94.1% when it is not used [10].

While most of the aforementioned researches have a close inclination to the research proposed in this paper, key aspects like privacy preservation of patient data, security of data, use of active learning as human-in the loop is hardly explored, however, they provide useful insights that are used to build our suggested approaches.

3 System Architecture

A description of the proposed architecture is given in this section, where the concentration is on the architecture components and the mode of operation.

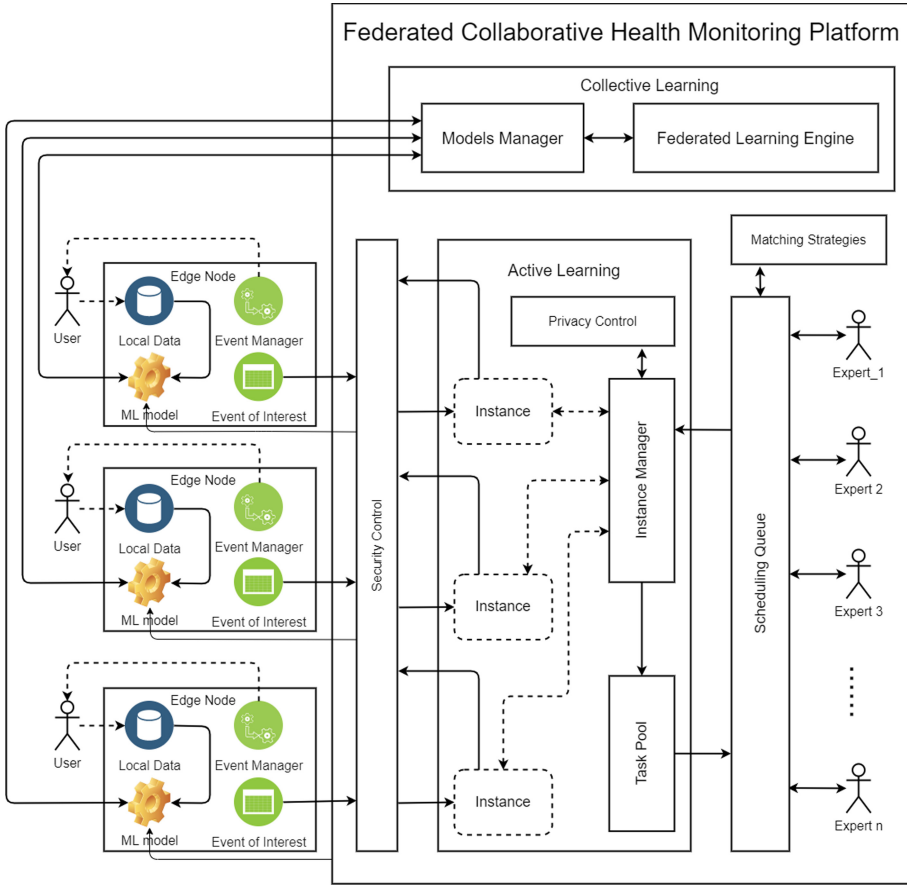


Fig. 1. The system architecture of federated collaborative health monitoring platform

3.1 Architecture Components

The proposed system architecture that is shown in Fig. 1 is composed of five different components that work in a coordinated mechanism in order to achieve the common objective and the role of each entity are shown as follows:

- **Collective Learning (Global Model):** The main role of the global model is to facilitate the aggregation of knowledge that originates from the edge nodes. For privacy concern, this platform will share only the learned knowledge that is coordinated by the collective learning component, while the personal or sensitive data are retained at the edge nodes. Since the nodes may have different types of data and ML models, the *Models Manager* is responsible for managing and coordinating the different models. The global model can be updated by way of synchronization between the edge nodes and the global model, which is managed by the *Federated Learning Engine*.

- **Edge nodes (Edge model):** Each received ML model from the collective learning component will be retrained over the personal local data. Then the model performs an incremental training approach in order to capture any new patterns or behavior. Moreover, the retrained edge models will be aggregated by the *collective learning* to re-update the global model's knowledge and redistribute it again to the edge nodes for further retraining. The edge nodes act as the clients of this platform who may belong to different stakeholders. For example, one node could be a hospital, which uses IoT devices to collect data from the patients and the hospital has its own local data and ML model to monitor the patients. When the hospital utilises this platform, it can increase the accuracy of its ML model by sharing learned knowledge with other stakeholders, and get emergent notification based on the opinions of the experts in the platform.
- **Security control:** Security control plays two major roles as follows: (1) Prevents, adversarial attacks, particularly, poisoning attacks, by creating a cryptographic hash during incremental training to retain the training data in its original form, and (2) Making it hard for a potential attacker to decipher the data contents as well as preventing malicious adversaries from accessing data when it is transmitted over the network to the experts during the Active learning process.
- **Active learning:** The edge nodes have their own strategies to send related data to the platform to get diagnosis from the experts. The strategies are managed by the *Event Manager* and the related data are sent via **Event of interest**. The validation responses from the experts will not only be used to give notifications to the related patients, but also gets annotated for the purpose of incremental learning. The instance in the active learning module that originates from any client is responsible for the related specific tasks from the client. The *Instance Manager* and *Privacy Control* need to pre-process the data, for example, by de-identification, before giving them to the experts. All the pre-processed data will be formatted and transferred to the *Task Pool*, which will assign the tasks to the experts. The expert validation responses, which prior also does checks with the client will be sent back to the *Instance Manager* where it can give the processed responses to the related clients.
- **Scheduling Queue:** Since there are many different types of experts with different levels, the tasks in the *Task Pool* are assigned to the appropriate expert based on the *Matching Strategies*. For example, the matching strategies could be based on the experts' reputation like specialty, experience of years and feedback in the past.

3.2 Modus Operandi

The suggested federated interactive IoT-based health monitoring platform that utilises domain experts (based on Fig. 1), is aimed at providing diagnose services to a variety of stakeholders, who own devices to monitor the users' health status. Consequently, each stakeholder can federate its ML model with the global model and create an instance in the active learning module in the platform to

get feedback from many experts provided by the platform. Each created instance in the *Active Learning* module is able to generate related tasks that are inclined to the corresponding expert based on the requirements. All the tasks are pre-processed to protect the users' privacy and pushed to the task pool, so they can be ready to be annotated by the experts. After the experts finish the tasks assigned to them, the results will be sent back to the *Instance Manager*, and the related instances will notify the respective stakeholders. Then the stakeholders are poised to take actions and update their ML models based on the feedback from the experts. Finally, the updated ML models (local nodes) in all the stakeholders can improve the global model via a federated learning architecture by way of transferring the extracted knowledge. Eventually, when the global model is updated, it can also synchronize with all the related stakeholders' ML models to improve their models' accuracy.

4 Experiment

This section details the experiment setting and performance analysis which is aimed at providing proof of concept of the proposition that has been mentioned in this paper. Furthermore, it is worth noting that the experiment's focus is to leverage federated learning to ensure data privacy is upheld while the ML models' accuracy is improved at run-time through continuous learning by relying on the multi-expert validations.

4.1 Experimental Setting

The study employs Continuous Ambient Sensors Dataset (CASA) human activity recognition dataset¹. The contents of CASA were collected from 30 different houses by using both ambient and PIR sensors. As listed in Table 1 each collected pattern comprises 37 features linked to different sensors distributed in those houses to monitor the user's daily activities. However, data linked to four houses were selected to conduct the experiment in order to evaluate the proposed architecture's behavior in terms of data privacy preservation by moving the ML model to the data location, and the improvement of the ML models accuracy's through the continuous 'human-in-the loop' learning.

To achieve this, Random-Forest classifier has been trained and tested over the selected data in both global and edge nodes. Hence, the selected houses data (csh105, csh108, csh111, and csh123) have been associated to the global model, node1, node2 and node3 respectively [11, 12]. At this stage, multiple local nodes (edge nodes) are trained over the local dataset's n number of iterations and the new learned knowledge is sent to the global node for aggregation.

¹ <http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+from+Continuous+Ambient+Sensor+Data>.

Table 1. CASA dataset features characteristics

Index	Features	Types	Index	Features	Types
1	lastSensorEventHours	Discrete	20	areaTransitions	Discrete
2	lastSensorEventSeconds	Continuous	21	numDistinctSensors	Discrete
3	lastSensorDayOfWeek	Discrete	22	sensorCount-Bathroom	Continuous
4	windowDuration	Continuous	23	sensorCount-Bedroom	Continuous
5	timeSinceLastSensorEvent	Continuous	24	sensorCount-Chair	Continuous
6	prevDominantSensor1	Discrete	25	sensorCount-DiningRoom	Continuous
7	prevDominantSensor2	Discrete	26	sensorElTime-Ignore	Continuous
8	lastSensorID	Discrete	27	sensorCount-Hall	Continuous
9	lastSensorLocation	Discrete	28	sensorElTime-Kitchen	Continuous
10	lastMotionLocation	Discrete	29	sensorElTime-LivingRoom	Continuous
11	complexity	Continuous	30	sensorElTime-Office	Continuous
12	activityChange	Continuous	31	sensorElTime-OutsideDoor	Continuous
13	sensorElTime-WorkArea	Continuous	32	sensorElTime-Hall	Continuous
14	sensorCount-Ignore	Continuous	33	sensorCount-Kitchen	Continuous
15	sensorCount-LivingRoom	Continuous	34	sensorCount-Office	Continuous
16	sensorCount-OutsideDoor	Continuous	35	sensorCount-WorkArea	Continuous
17	sensorElTime-Bathroom	Continuous	36	sensorElTime-Bedroom	Continuous
18	sensorElTime-Chair	Continuous	37	activity	Text (class label)
19	sensorElTime-DiningRoom	Continuous			

4.2 Performance Analysis

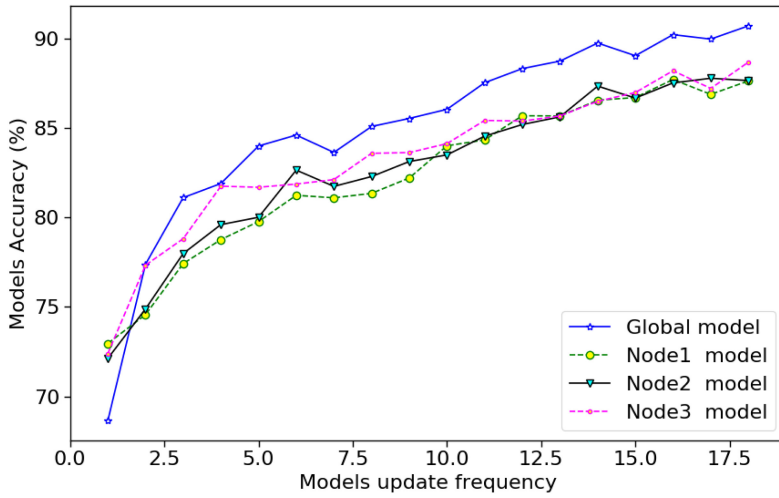
In our performance analysis, we evaluate the ultimate accuracy of the global model after aggregating the knowledge from the local edge nodes (see Fig. 2a and Fig. 2b). Observations from this analysis are presented as follows:

- In the first iteration the global and the edge nodes exhibited low accuracy.
- After aggregating the trained edge (distributed) models by taking the average of the learned knowledge, the model’s learning behavior improved tremendously.
- The learned knowledge has been used in the derivation of the new (current) global model, which is then redistributed to the linked edge nodes for further training.
- Interactive learning has successfully aided the edge ML model’s incremental learning by utilising the new annotated data. This data has played a significant role in the improvement of the model’s learning process.

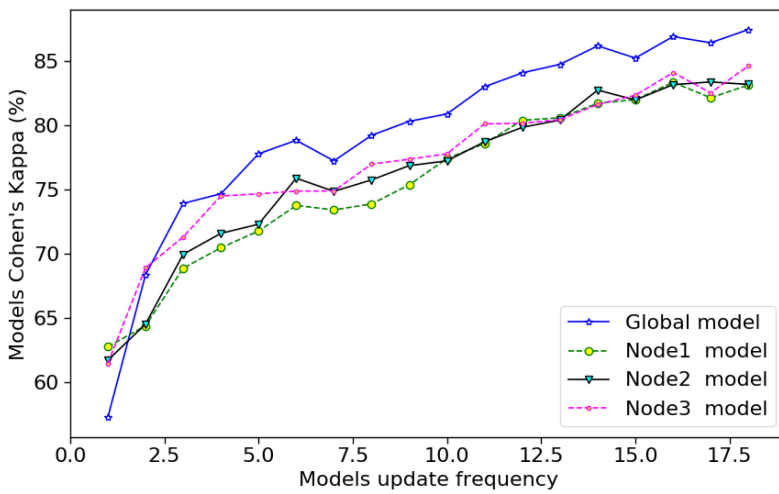
From the Figs. 2a and 2b, we notice that the performance of the global model is basically influenced by the availability of the linked edge nodes and the user-feedback from the experts which has shown an improvement on the model accuracy and performance.

4.3 Security and Privacy Analysis

In our security and privacy analysis, we hypothesise the threat model, a variety of possible attacks on the proposed IoT-Health platform and then an analysis is given on the same.



(a) Both federated learning model and participants nodes models accuracy



(b) Both federated learning model and participants nodes models Cohen's kappa

Fig. 2. Experiments depicting the federated global model and the participants of local nodes

Threat Model. Our threat model makes the following assumptions on adversarial perspectives: In the context of IoT-Health platform, there may exist malicious content that may hinder the global and local models accuracy, and this content may be channelled to the training data in an adversarial training attack at the local nodes during incremental learning. Also, we assume that an attacker may actively defeat the security of the platform by challenging the ciphertext in order to eavesdrop all communication when data is sent over the network.

Attack Analysis. Based on the assumptions of the threat model (Sect. 4.3), we analyse the security attacks as follows:

- Poisoning attack: An adversary may subject the training data to malicious content, which may end up affecting the knowledge that is extracted from the local nodes, thus affecting the accuracy of the global model: We have suggested the use of cryptographic hashes to maintain the integrity of training data at the local nodes.
- Ciphertext attack: Based on the security control technique in Fig. 1, we assume the role of an adversary is to corrupt sensitive patient data. Therefore, an adversary may obtain encrypted data or the secret keys to have a direct access to the data. We suggest the use of strong encryption approaches like homomorphic encryption and differential privacy.
- Eavesdropping attack: An adversary can attempt to eavesdrop on data that is sent between the local node and the global model, which ultimately has an impact on data privacy. We suggest maintaining strong privacy techniques which are discussed next.

Privacy Analysis. In our approach towards privacy-preserving technique, we have proposed an approach that utilises, federated learning and concepts of interactive learning that collaboratively are able to build a global model without sharing data whatsoever. Data is retained at the local nodes, where each node is able to maintain its data (Fig. 1). From this, the new global model is only able to learn from the knowledge from the local nodes which ensured that privacy is preserved locally, owing to the fact that only knowledge is transferred to the global model. Nonetheless, personal data of users for which the system was trained on might still be revealed indirectly through privacy attacks on the machine learning model. In particular, we identify the following attacks:

- Model inversion attack: An adversary having access to some data belonging to specific patients included in the training data, can infer further data about those same individuals by observing inputs/outputs of the machine learning model. For example, given some demographic information an adversary could infer genetic markers from the model despite having only partial access to the underlying training data [13]. A common mitigation against model inversion attacks is differential privacy.

Table 2. Evaluation of identified security and privacy issues

Security & privacy issue	Overview	Mitigation approaches
Poisoning attack	Contaminating training data with malicious content. The ultimate knowledge from the training is falsified. Affects the accuracy of the model	To retain the integrity, cryptographic hashes are preferred given that they are deterministic, where same input guarantees same output the fact that they are irreversible
Ciphertext attack	It is possible to tamper with encrypted data at the local nodes, by obtaining the public key used to encrypt the data from a source	Not only using strong encryption but employing digital certificate during data transmission
Eavesdropping attack	An adversary or a malevolent data labeler can listen to or gain access to data being transmitted with an elevated privilege	During incremental training and during the provision of learning model updates to use blockchain for it guarantees secure data transmission
Model inversion attack	An adversary making an inference about the data in possession	Use of differential privacy
Membership inference attack	Trained models can be used to leak information about a patient's record	Use of regularization
Model stealing attack	Constructing surrogate models from extracted model parameters	Use of information laundering

- Membership inference attack: An adversary may deduce whether a given patient is present in the training data of a machine learning model. For instance, if hospital records are used to train a model which predicts when a patient will be discharged from the hospital, adversaries could use that model with other data to reveal whether an individual had visited one of the hospitals that generated the training data during the period the data was collected. The use of regularization, e.g., through L_2 regularization, is identified as a technique for reducing membership inference attacks.
- Model stealing attack: Adversaries may extract parameters from a target model allowing them to reconstruct a surrogate model with similar performance as the target model. While this attack is harder to conduct in a federated learning setup as is the case for FIL-IoT-HMP, where multiple decentralized edge nodes are involved, in theory this attack may still be possible. Model stealing may indirectly compromise privacy, but more so, the confidentiality of the health platform users. Information laundering is a technique that can be used to mitigate against model stealing attacks.

Based on how the FIL-IoT-HMP model has been positioned, we argue that blockchain technology has been presented as a more suitable technique for enhancing secure data sharing at the local nodes by providing tamper-free adversarial attacks during incremental training [4]. Normally, adversarial attacks during active learning are common occurrences based on existing learning threat

landscape like targeted attacks, unusual propagation attacks, malicious logic insertion and overall system manipulation [14].

Given that federated learning model is shared across the multiple nodes, we also argue that the following aspects transpire as a result, however a summary is given in Table 2:

- The data from the IoT health platform is regarded to be sensitive, as a result privacy preservation is a key aspect of consideration in this context
- At the edge, blockchain integration gives an assurance of the following: Secure data sharing during incremental learning process, resource location where smart contracts can be used as a way of access control and management
- From a security perspective, existing vulnerabilities arising from the learning model, especially during data transmission may enable an attacker to launch specific attacks that can lead to leakage of sensitive information. In this perspective federated learning guarantees privacy protection and verification through periodic updates during the transmission of learning models.

5 Conclusion

We have proposed a federated interactive IoT-based health monitoring platform that utilizes (active, interactive and human-in-the loop) This platform has a strong privacy-preserving feature and also its able to counter adversarial attacks during incremental learning. The problem of data leakage has been analyzed correctly by allowing the global model to only share the knowledge from local nodes while the data is retained at the local nodes. For future work, we aim to extend this work to incorporate multiple machine learning algorithms using different datasets in order to study the effect of the expert validation.

References

1. Kumar, P.M., Gandhi, U.D.: A novel three-tier internet of things architecture with machine learning algorithm for early detection of heart diseases. *Comput. Electr. Eng.* **65**, 222–235 (2018)
2. Ponemon Institute. 2020 State of Password and Authentication Security Behaviors Report. <https://pages.yubico.com/2020-password-and-authentication-report/>
3. Montalbano, E.: Kaspersky: Attacks on Smart Devices Rise Threefold in 2018. <https://securityledger.com/2018/09/kaspersky-attacks-on-smart-devices-rise-threefold-2018/>. Accessed 29 Mar 2021
4. Kebande, V.R., Alawadi, S., Bugeja, J., Persson, J.A., Olsson, C.M.: Leveraging federated learning & blockchain to counter adversarial attacks in incremental learning. In: 10th International Conference on the Internet of Things Companion, pp. 1–5 (2020)
5. Yang, Q., Liu, Y., Chen, T., Tong, Y.: Federated machine learning: concept and applications. *ACM Trans. Intell. Syst. Technol. (TIST)* **10**(2), 1–19 (2019)
6. Holzinger, A.: Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Inf.* **3**(2), 119–131 (2016)

7. Yuan, A., Yan, L., Cai-Wen, M., Li-Min, S., Zhi-Feng, X.: SENS-U: remote human in loop health-monitoring system at home. In: 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, vol. 1, pp. 441–445. IEEE (2008)
8. Saeedi, R., Sasani, K., Gebremedhin, A.H.: Co-meal: cost-optimal multi-expert active learning architecture for mobile health monitoring. In: Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, pp. 432–441 (2017)
9. Saeedi, R., Sasani, K., Gebremedhin, A.H.: Collaborative multi-expert active learning for mobile health monitoring: architecture, algorithms, and evaluation. *Sensors* **20**(7), 1932 (2020)
10. Chen, Y., Qin, X., Wang, J., Yu, C., Gao, W.: FedHealth: a federated transfer learning framework for wearable healthcare. *IEEE Intell. Syst.* (2020)
11. Alawadi, S., Delgado, M.F., Pérez, D.M.: Machine learning algorithms for pattern visualization in classification tasks and for automatic indoor temperature prediction. Ph.D. thesis, Universidade de Santiago de Compostela (2018)
12. Alkhabbas, F., Alawadi, S., Spalazzese, R., Davidsson, P.: Activity recognition and user preference learning for automated configuration of IoT environments. In: Proceedings of the 10th International Conference on the Internet of Things, pp. 1–8 (2020)
13. Veale, M., Binns, R., Edwards, L.: Algorithms that remember: model inversion attacks and data protection law. *Philos. Trans. Roy. Soc. A: Math. Phys. Eng. Sci.* **376**(2133), 20180083 (2018)
14. Kebande, V.R., Alawadi, S., Awaysheh, F.M., Persson, J.A.: Active machine learning adversarial attack detection in the user feedback process. *IEEE Access* **9**, 36908–36923 (2021)