Computers in Biology and Medicine 167 (2023) 107630

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

A blockchain-based federated learning mechanism for privacy preservation of healthcare IoT data

Wided Moulahi [a,b], Imen Jdey [a,b,*], Tarek Moulahi [c], Moatsum Alawida [d], Abdulatif Alabdulatif [e]

[a] Faculty of sciences and Techniques of Sidi Bouzid, University of Kairouan, Tunisia
[b] REsearch Groups in Intelligent Machines (LR11ES48), Tunisia
[c] Department of Information Technology, College of Computer, Qassim University, Kingdom of Saudi Arabia
[d] Department of Computer Sciences and Information Technology, Abu Dhabi University, 59911, Abu Dhabi, United Arab Emirates
[e] Department of Computer science, College of Computer, Qassim University, Buraydah, Kingdom of Saudi Arabia

ARTICLE INFO

Dataset link: https://kaggle.com/datasets/uciml/pima-indians-diabetes-database

Keywords:
Blockchain
Federated Learning
Internet of Things
Healthcare IoT
Machine Learning
Privacy preservation
Smart Contract

ABSTRACT

The Corona virus outbreak sped up the process of digitalizing healthcare. The ubiquity of IoT devices in healthcare has thrust the Healthcare Internet of Things (HIoT) to the forefront as a viable answer to the shortage of healthcare professionals. However, the medical field's ability to utilize this technology may be constrained by rules governing the sharing of data and privacy issues. Furthermore, endangering human life is what happens when a medical machine learning system is tricked or hacked. As a result, robust protections against cyberattacks are essential in the medical sector. This research uses two technologies, namely federated learning and blockchain, to solve these problems. The ultimate goal is to construct a trusted federated learning system on the blockchain that can predict people who are at risk for developing diabetes. The study's findings were deemed satisfactory as it achieved a multilayer perceptron accuracy of 97.11% and an average federated learning accuracy of 93.95%.

1. Introduction

Since its formal proposal in 1956 [1], Artificial Intelligence (AI) has been steadily and rapidly proliferating itself in everyday life. It comes to consolidate human efforts to face challenging issues and solve problems of humanity. As it is expanding its spheres into many different areas, AI is indeed a crucial support system for the development of society.

Taking advantage of the vast amount of data produced across the world and the significant advancements in hardware technology, the field of Machine Learning (ML) has advanced to the point where a number of intelligent systems built on ML models are beginning to find their way into mainstream applications leveraged in performing day-to-day human tasks [2]. This progress is extremely beneficial and valuable since it enhances people achieve higher efficiency and productivity. There is a wide range of uses for machine learning, from simple applications like recommender systems [3] used by information-based companies like Google, Twitter, Netflix, and LinkedIn to sophisticated applications like self-driving cars [4], smart healthcare [5], earthquake prediction [6], and many more [7]. One sector that has greatly benefited from these technological advancements is the healthcare sector. Machine learning has shown remarkable applicability in the medical field, revolutionizing healthcare practices and patient outcomes [8,9]. More importantly, smart healthcare systems have opened up virtually unbelievable prospects and novel possibilities for saving lives. Their potential lies in their ability to analyze vast amounts of patient data, identify patterns, and generate insights that aid in accurate diagnoses and treatment planning. HIoT enables a tremendous change in the field of digital health, as evidenced by the findings of the study by Habib Zadeh and colleagues from 2019 [10]. By deploying a wide range of intelligent systems, it aspires to offer consistent medical attention to patients. Thus, it is an essential part of keeping people alive. Research interest in the domain of HIoT is therefore growing at an exponential rate.

However, being impressed by the humongous opportunities that AI offers for a smarter life should not lead to undermining the serious risks that it poses. The need to shed light and get insights into the associated risks with AI and ML is emerging. In fact, patient confidentiality is of utmost importance, and the aggregation of sensitive medical information raises concerns about data breaches, unauthorized access, and potential misuse. Medical data contains highly sensitive and personal information about patients, including their medical history, diagnoses, treatments, and genetic information. Any unauthorized access to or disclosure of this data can lead to severe consequences

* Corresponding author at: Faculty of sciences and Techniques of Sidi Bouzid, University of Kairouan, Tunisia.
E-mail address: imen.jdey@fstsbz.u-kairouan.tn (I. Jdey).

https://doi.org/10.1016/j.compbiomed.2023.107630
Received 8 June 2023; Received in revised form 27 September 2023; Accepted 23 October 2023
Available online 31 October 2023

for the individuals involved, including identity theft, discrimination, or social stigmatization. The medical field could be a prime target for cyberattacks due to the value of medical data on the black market. Privacy breaches resulting from cyberattacks can lead to significant harm to individuals and organizations.

Though ML models have been established to outperform well across a diverse range of tasks, they are also susceptible to adversarial attacks [11]. The latter category attempts to trick an ML model into failing in a variety of ways, leading to misclassified results. The attacker in a black-box attack does not have access to or knowledge of the model's internal settings. A white-box exploit, on the other hand, allows the attacker complete access to all model parameters, allowing for significant alterations to the model's architecture and, by extension, its outcomes. If the health-related ML model is tricked, it will lead to misclassified results that put lives at risk.

Given the aforementioned reservations, academics, researchers, and stakeholders are battling with the difficult challenge of how to develop a trusted AI [12]. While using IT solutions in healthcare saves lives, it also raises a number of concerns, particularly in the areas of privacy and security. The confidentiality, integrity, and availability of information must be safeguarded to prevent unauthorized access, use, disclosure, interruption, alteration, or destruction, as defined by [13]. Nevertheless, privacy safeguards ensure that users' information is handled in line with their preferences when it is collected, stored, utilized, and disseminated. Data privacy means being in charge of one's own information. However, security is tied to safety. It is the process of protecting data against various vulnerabilities. Security is often compared to a game of chess, and as such, it is important to not only develop a successful strategy but also to anticipate potential threats from the adversary and their possible responses.

Striking a balance between leveraging the power of machine learning while ensuring stringent privacy protection measures is crucial to foster trust and encourage further advancements in the medical domain. Therefore, researchers are always working on finding new approaches to reinforce ML's robustness in the face of the ongoing conflict between ML and attack vectors. They take advantage of the state of the art and cutting-edge technologies to develop formidable defenses. Some of the most promising alternatives include federated learning (FL) and blockchain technology. FL, as an emerging shift in ML paradigms, comes with promises to preserve privacy. Instead of sending data to a model, this approach reverses the process and brings the model to the data. The blockchain is yet another effective new innovation. Its robust features, ensuring both content and operations security, make it a potential solution for combating malicious attacks. Driven by privacy and security concerns, this research tries to consolidate efforts for a more secure and efficient smart healthcare [14]. The primary goal is to construct a secure and private decentralized learning platform using blockchain technology-concisely, a blockchain-based federated learning system that ensures privacy preservation and data integrity.

1.1. Research contributions of this work

In order to reach the main goal of this study, two main objectives have been set:

1. To build an ML system based on FL in order to guarantee privacy preservation.
2. To integrate blockchain technology into the FL system to ensure ML model integrity and immunize the system against attacks.

This paper is organized as follows: Section 2 provides the background of the study. Section 3 outlines an overview of the related work. The proposed framework to solve the addressed problem is presented in Section 4. The experiments and the results are presented in Section 5. The achieved results are discussed in Section 6. Section 7 presents the threats to validity. The conclusion is given in the last section of the paper.

2. Background study

This section is subdivided into three subsections. First, we present an overview on the HIoT, secondly the FL and finally the Blockchain technology.

2.1. Healthcare internet of things (HIoT)

The World Human Organization (WHO) Constitution envisages "*the highest attainable standard of health as a fundamental right of every human being.*"[1] With an ageing global population, rising life expectancy, and the emergence of new diseases, it is imperative that healthcare evolve to be more proactive rather than reactive – that is – preventive than curative.

Health has a significant impact on a country's economic development. People who live longer and healthier lives tend to contribute more to society as a whole. As a result, the healthcare sector is consistently innovating in order to fulfill this fundamental human right.

The healthcare sector has always benefited from technological advancements. This helped bring forward a novel outlook to medical care. Now, instead of patients going to get care, it is the healthcare providers that go to the patients. Currently, we talk of electronic health (e-Health) [15].

Internet of Things (IoT) is one of the main technologies enabling sustainable healthcare delivery, which is important since putting healthcare services out of reach of regular people puts their lives at risk. IoT has slowly proliferated itself in healthcare. However, pre-Pandemic conditions have altered a lot of situations following the COVID-19 pandemic. The need for continuously available (always-on) healthcare services and the social distancing led to a dramatic transition in IoT deployment in healthcare. The technology was hugely applicable to tackle the pandemic. Combating the epidemic would benefit greatly from this technology. In many health systems, for instance, the global number of online consults has increased by a factor of 50 to 100 due to the usage of telehealth and telemedicine to improve healthcare solutions and to remotely treat patients [16,17]. In light of the recent coronavirus outbreak, Dr. Peter Gocke, Chief Digital Officer at Charite University Medical Center in Berlin, has stated that "*we need more open and transparent communication to change structures and processes that are no longer working*".

In the following three subsections, we will discuss typical uses for HIoT, security and privacy concerns with HIoT deployment, and security architecture to solve these issues.

2.1.1. HIoT applications

IoT devices help in patients' surveillance [18]. In nursing homes and hospitals, they help keep tabs on residents' health. They help disabled persons who live alone by keeping an eye on vitals like heart rate and blood pressure, among other things. Medical Fridges, Fall Detection, Dental are other applications of IoT based healthcare systems [19]. In 2019, 86% of healthcare providers used IoT in some capacity. Forbes predicts that by 2020, 646 million IoT devices will be in use in healthcare facilities such as hospitals, clinics, and doctors' offices [20]. HIoT can be used in a variety of contexts. Here we highlight some of the most popular uses of HIoT [10,19]:

- **Blood Glucose Monitoring:** A diabetic patient needs to keep a close eye on their blood glucose levels. It provides advice on how to cope with the illness. Warnings and useful recommendations can be issued by IoT devices in potentially harmful Warnings and useful recommendations can be issued by IoT devices in potentially harmful

[1] https://www.who.int/news-room/fact-sheets/detail/human-rights-and-health

Table 1
Cloud threats and their descriptions.

Threat	Description
Malicious Insider	Someone from the inside having access to user's data may manipulate it.
Data Loss	A malicious user with an unauthorized access may alter or delete the existing data.
Man-in-the-middle (MITM)	A kind of Account Hijacking threat. The attacker can alter or intercept messages in the communication between two parties.
Cloud Storage Data Exfiltration	Sensitive data is viewed, used or stolen by someone outside of the organization's environment.
GANs-based Inference Attacks	Generative Adversarial Network-based attacks launch inference and poisoning attacks.
Communication bottlenecks	The communication bandwidth gets weaker and drastically disrupts the FL process.
Backdoor attack	Injecting malicious task into the current model without affecting its actual accuracy.
Organized crime and hackers	Advanced persistent threat (APT) groups target thieving to data acquisition.

- **Cardiac Monitoring:** The mortality rate from cardiovascular disease can be reduced with the use of continuous cardiac monitoring, according to a recent study cited in the journal Zagan 2017 Healthcare [21]. IoT devices aid in monitoring of hypertension and hypotension [22].
- **Alzheimer's Disease (AD) Monitoring:** An AD patient needs round-the-clock care. Using IoT devices, caregivers can perform things like track a patient's whereabouts and receive alerts about unusual behavior [23].
- **Medical Fridges:** IoT systems are used to control internal conditions in freezers where medicines, vaccines and organic elements are stored [24].
- **Fall Detection:** Detecting falls is a major concern for the elderly and the crippled who live alone, and here is where IoT solutions shine [25].
- **Dentistry:** Internet of Things is also used in dentistry. Bluetooth connected toothbrush analyzes the brushing uses and gives information on the brushing habits for personal use or for dentists [26].

2.1.2. Security and privacy challenges

HIoT has greatly simplified our daily life. However, widespread adoption is unlikely because of concerns over the security and privacy of user data. In fact, people almost feel uncomfortable with sharing personal medical data. These concerns are raised with HIoT devices' susceptibility to attacks. Some of these security issues are as follows [27]:

- **Unauthorized Access to Radio Frequency Identification (RFID):** Unauthorized RFID access is a key problem in the Internet of Things. If someone has access to the user's identification tag, it could compromise their privacy. The tag can be read, changed, and possibly even broken [27].
- **Sensor-Nodes Security Breach:** A sensor node's primary function is to act as a data collector and transmitter. As part of a two-way sensor network, it can sometimes collect and relay information. It is also possible to get data. That could make the node vulnerable to outside attackers. Intruders can either steal or tamper with node data. The attacks of jamming, tampering, Sybil, and flooding, among others, are summarized by [28].
- **Cloud Computing Abuse:** The cloud-centric architecture exposes IoT systems to several threats. For example, if an attacker gets access to the server, it can upload any malicious software and get control on the connected devices. Table 1 summarizes some of these serious threats [27,29,30].

2.1.3. Security architecture

Regarding the previously mentioned challenges, establishing a security architecture becomes an increasing need to ensure HIoT efficiency. Several actions could be taken in different levels to face such challenges. In general, security architecture contains three levels [20] as presents Fig. 1:

- **Device security (perception layer):** At this level, the threat is physical. It attacks the physical devices. The protection should then target the device itself.
- **Communication securities (Network and transport layer):** It aims to protect network and communication.
- **Cloud security (Application layer):** It aims to protect the data itself against threats of different natures.

2.2. Federated learning

Coined by Google in 2016 and referred to as Vanilla FL, the FL aims to generate ML models from distributed datasets on various devices. It came with promises to protect data privacy. And since, it is experiencing a great growth in both academia and industry. Privacy and security are among the essential objectives of FL. FL allows private data to remain under the control of its owner. It is built on the idea of bringing the model to the data, but it goes beyond that. It allows benefiting from peers' experiences. In a FL system, there are two principal actors [31,32] as below:

- **Server:** it orchestrates the training process and updates the learning model without getting access to the client data.
- **Client:** each client holds its own data, trains the model locally and shares the updates with the others via server.

The number of clients participating in the learning process is variable. However, there is usually only one server.

The following parts present the FL life cycle and the challenges it faces.

2.2.1. Federated learning life cycle

From the baseline model to the final model, the FL life cycle goes through several steps. According to [33], a FL process is composed of continuous communication rounds between the server and the end devices. It is completed once the desired accuracy or the fixed number of rounds are reached.

First, the server generates a generic model then each round follows these steps as presented in Fig. 2:

1. The server selects a subset of clients and sends them the generic model.

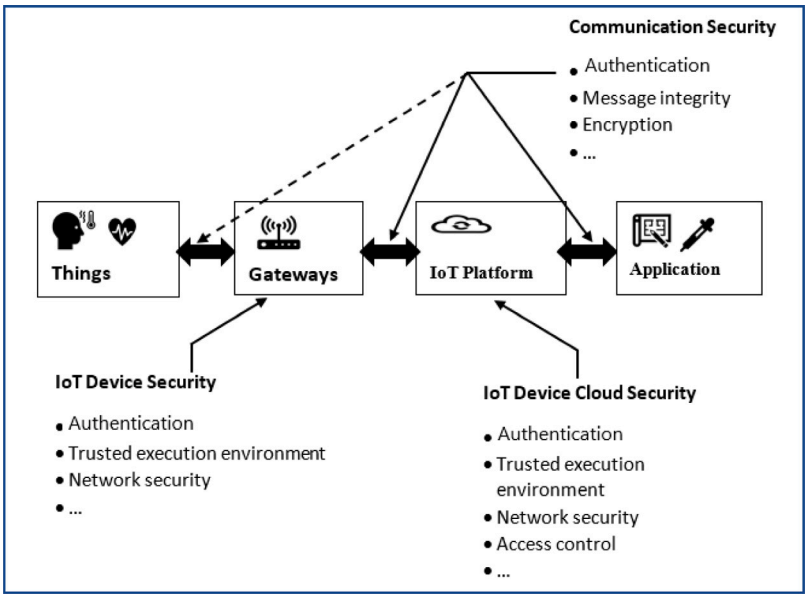

Fig. 1. Security architecture: Levels and descriptions.

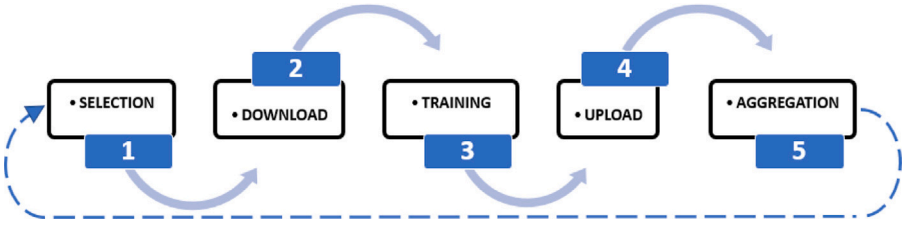

Fig. 2. Federated learning life cycle, continuous communication rounds.

2. The selected clients download the current model parameters and initialize the local model with such parameters.
3. Each client trains the model on its own local data. At the end of the training, it updates and optimizes the model parameters.
4. The clients send the optimized new parameters to the server.
5. Based on an aggregation algorithm, the server aggregates the clients' updates. The global model is updated with the new aggregated parameters.

The updated model is sent back to the selected clients. A new round takes place until getting the final model [34].

2.2.2. Challenges

FL's primary guarantee is the protection of user privacy. Consequently, impressive results are being shown in that setting. On the other hand, it has a number of problems that need addressing [33].

- **Non-Independent and Non-Identically Distributed (Non-IID) Data:** Each customer base is unique, and as such, each client has its own experience. The generated dataset is, therefore, dependent on its personal behavior. The same device may be used by different people. That makes the dataset not representative and even non-identically distributed from one device to another [35].
- **Unbalanced Data:** Dataset sizes might vary widely from one device to another because of the diverse range of ways in which devices are put to use. The effectiveness of the outcomes is reliant on the quality of the training dataset [28]. Thus, it is important to consider the size of each client's dataset when performing an efficient aggregate.
- **Massively Distributed Data:** Smartphones, Internet of Things (IoT) devices, automobiles, institutions, and many others are all examples of the kinds of clients that can take part in a FL process. Due to this, there is a great deal of variation in the statistics, leading to highly distributed data [33].
- **Unreliable Device Connection:** A client's accessibility is tied to its network connectivity. If a client has a poor or slow connection, they may be unable to take part in the communication round. There could be a drop in the performance of the model as a result of that, affecting the model's efficiency [33].
- **Limited Device Memory:** Devices differ in their computational capacities and amounts of available memory. It is likely that a smartphone or IoT device has less storage space than other gadgets [33].
- **Single point failure:** FL depends on a centralized server to aggregate model updates. This aggregator could be the target of an attack. In addition, a volume of data transmitted to the server simultaneously may overwhelm its capabilities [36].
- **Privacy leakage:**. Substantial information may be gleaned from intermediate updates via inference attacks. It is also possible for a hostile server to use a Generative Adversarial Network (GAN) to steal private information [37].
- **Lack of motivation:**In order for FL to work, it is assumed that all nearby devices are willing to take part in the training process and share their experience. Unfortunately, that cannot be accomplished. Some nodes may choose not to take part in the poll for various reasons, such as a preference for privacy, a dearth of computing resources, or even just plain selfishness etc [33].
- **Poisoning Attacks:** FL guarantees the safety, security, preservation, and integrity of your data. However, attacks involving poison are always a concern. They pose a risk to system security and can originate either from the server or the client side [38].

Different techniques were associated to FL to deal with privacy issues: – Homomorphic encryption: it is an encryption technique that allows computing encrypted data, without the need to decrypt it. Data could be used without revealing its content, and even the results of its manipulation remain encrypted [39,40]. – Secure aggregate (SecAgg): it consists to encrypt the model updates from each participating device before sending them to the central server. The server then uses a "secure multiparty computation" technique to aggregate the encrypted updates without decrypting them [30,39,41]. – Differential Privacy: it consists of injecting a carefully chosen amount of noise into the data without loosing its utility. That makes it hard to infer and, at the same time, ensures getting significant results [39].

One other candidate solution to address these privacy concerns and other FL issues such as single point failure is Blockchain technology [42], which can help users maximize the advantages of FL. In the next section, we will dive deeper into blockchain.

2.3. Blockchain

There is no single, comprehensive description of what blockchain actually is. There are other comparable definitions in the literature. Thus, the literature provides various but similar definitions. The term "Blockchain" is self-explanatory. It consists of blocks connected to each other. Blockchain allows data storage using consensus algorithms to generate and update the distributed ledger, and encryption to ensure security during transmission [43].

Blockchain, as defined by [44], is a decentralized ledger that records transactions between users independently. It is a distributed database without a central authority. It manages the flow of information between nodes in a peer-to-peer network. It is impossible to delete information once it has been transferred. Data transfer is irreversible. The following subsections outline the Blockchain characteristics, the components leading to these firm characteristics, and two vital technologies that Blockchain relies on and owes its growth to.

2.3.1. Blockchain characteristics

Blockchain robust features make it an attractive option for protecting the integrity and privacy of sensitive information [45]:

- **Decentralization:** Blockchain operates as a distributed ledger system. There is no central authority to control transactions. Unlike a centralized architecture, a transaction can be conducted between any two nodes without the intervention of a third, trusted party. That helps to avoid the single point failure problem typical of server-based architecture [45].
- **Transparency:** Blockchain transactions are broadcast to all participating nodes. Thus, all participating nodes can freely access them. Any node can have an overview on the recorded transactions [45].
- **Immutability and Traceability:** Once a transaction is validated and stored in the ledger, it cannot be modified unless a node has control on over 51% of the nodes. To trace previous records, it just needs to access any node in the network and consult the associated ledger [45].
- **De-Trusting:** In Blockchain, a user's identity need not be revealed in order to connect with other nodes. The nodes in a peer-to-peer (P2P) network are not required to trust each other to validate transactions. Instead, the trust comes from the adopted consensus algorithms. The latter control the communication between nodes and ensure that the distributed ledgers are the same [45].
- **Anonymity:** A user can interact with other nodes in Blockchain without revealing its identity. It is identified by a randomly generated address [45].

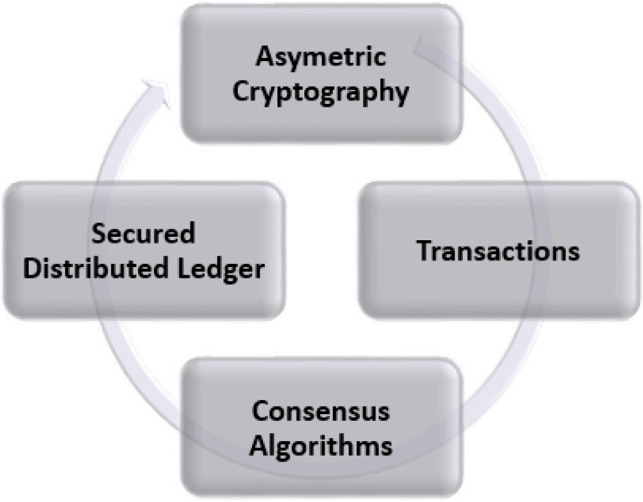

Fig. 3. Blockchain components and their interactions.

2.3.2. Blockchain components

Blockchain owes its firm characteristics to its different components which interact to ensure a secure data sharing as Fig. 3 shows.

- **Transactions:** Blockchain enables sharing information between the different nodes among a P2P network. The source node generates a file containing the exchange information. Then, it broadcasts it to the entire network for validation. A set of validated transactions is congregated in blocks. The first generated block is called Genesis Block. Each transaction changes the Blockchain state. Each node holds a log containing the history of previous transactions.
- **Secured Distributed Ledger:** The history of validated transactions is recorded in a Ledger. A block generated is broadcasted for validation. Once a block is validated, it is queued to the precedent blocks forming, then, a chain [46].
- **Asymmetric key Cryptography:** To achieve security, Blockchain relies on two types of cryptography and Hash Function:
  Hash Function: A hash function is a cryptographic implementation of a hash algorithm. It is a way to apply a hash algorithm to an input. Its result is known as a message digest. It has a unique fixed-size. Blockchain uses the Secure Hash Algorithm (SHA). Its size is 256 bits (SHA-256). This is a one-way hashing function. In most cases, the digest message cannot be used to reconstruct the original message. Finding an input that provides a certain message digest is similarly challenging. Two distinct inputs will never provide the same hash value [47,48].
  Asymmetric Encryption: Public-key encryption is another name for asymmetric encryption. There are two keys needed to encrypt and decrypt any message. It has two main applications: addresses and electronic signatures. A node must have a wallet inside the P2P network to conduct any transaction. This wallet is protected by the private key, while its address uses the public key [47,49].
- **Consensus Mechanism:** Unreliable nodes in a P2P network require a consensus method. Thus, the P2P distributed network necessitates a mechanism to reach consensus between the untrustworthy nodes. Several consensus algorithms are used to control communication and ensure that ledgers in different nodes are consistent: Proof of Work (PoW), Proof of Stake (PoS), Practical Byzantine Fault Tolerance (PBFT), Delegated Proof of Stake (DPoS) etc [45,50]. Table 2 summarizes the characteristics of each algorithm.

Table 2
Consensus algorithms characteristics and usage examples.

	PoW	PoS	PoET	PBFT
Blockchain Type	Public	Public	Private	Private
Computational Problem Solver	Miner	Validator	Miner	Generator
Energy Consumption	High	Less	Less	Less
Efficiency	Less	Less	High	Less
Usage Example	Bitcoin	Ethereum	Sawtooth Hyperledger	Fabric Hyperledger

2.3.3. Blockchain technologies

The Blockchain growth is mainly due to the advanced technologies it deploys. Smart Contracts and Hyperledger are two vital technologies adopted by Blockchain.

- **Smart Contract :** The term "smart contract", or "SC", was coined by Nich Szabo in 1994. A SC is a code block that executes itself when certain conditions are satisfied [51]. Launched in 2015, it is the first of its kind to be used in the Ethereum Blockchain. It is difficult to make adjustments to a SC after it has been put into production. SCs are crucial for establishing connections — basically setting and getting data to and from the distributed ledger [43]. In spite of their promising appearance in the future, SCs are not yet capable of functioning independently as a decentralized server.
- **Hyper Ledger :** Hyperledger is an umbrella term for a collection of Linux Foundation-sponsored open-source initiatives. These projects are used for Blockchain development. Hyperledger differs from permissionless Blockchain systems like Ethereum since it is generally private and permissionless. It serves as a "greenhouse" that promotes collaboration between businesses and developers [52].

3. Related work

Blockchain-based federated learning (BCFL) has been employed in several fields. Industrial IoT, medicine, Internet of vehicles are some of BCFL applications. The different applications of BCFL in medicine are presented as follows.

To face covid-19 crisis, the work of [53] used a Blockchain managed FL approach. It allows detecting the Corona virus from symptoms. The proposed approach relies on two modules: An immunization model based on Deep Learning algorithms [54,55]. It detects covid-19 from symptoms data such as loss of smell and taste, fever, vomiting, cough, shortness of breath. A biometric module using neural network architecture is used for face recognition. A set of HIoT sensors are used to collect data [56]. The two modules coupled with an AI-based QR code generator produce a Blockchain mapped health immunization certificate. The system that allows a dynamic health QR code visualization has applications for three actors: the patient, the health authority and the stakeholders wherever health status must be checked for safety. The transactions through the whole system are stored in Blockchain. The private raw data is stored off-chain in the InterPlanetary File System (IPFS) repository. The training data hash is then stored in the Blockchain. Two types of Blockchain were deployed. Ethereum and Hyperledger were used to manage the distributed ledger system. The system added security and privacy algorithm to FL ecosystem. It aims to find a balance between the model accuracy and the privacy budget.

The authors of [57] proposed a method to help medical practitioners to detect Covid-19 using lung Computed Tomography (CT) scans. The architecture relied on a Modified Capsule Network for Covid-19 subjects' classification. The dataset contained 34,006 CT images belonging to 89 persons, 68 among them were diagnosed as positive and the other 21 were diagnosed as negative cases.

In [58], the authors proposed a BCFL approach to predict diabetes. An Artificial Neural Network model was trained on Pima Indians Diabetes Database provided by the American National Institute of Diabetes and Digestive and Kidney Diseases. The dataset was distributed among 15 participants in a randomized way to simulate a federation scenario. SCs in an Ethereum Blockchain were used instead of a centralized aggregator. They collect model updates (weights and biases) from the nodes participating in the training and send them the aggregated parameters to update their models. The model was trained in different circumstances. The average accuracy of the model is 73% (against 76% in a centralized architecture). The developed model is partially secure. Inference attacks are still possible. Privacy preserving protocols can be used to improve the model performance.

The work of [59] proposed a BCFL approach for smart healthcare based on Medical Internet of Things (MIoT). It aims to resist to the single point failure problem, to protect data privacy and detect poisoning attacks. A Convolutional Neural Network (CNN) was trained on Pima Indians Diabetes Database to generate a diabetic prediction model. A small Blockchain prototype based on Ethereum was used to store, verify and share the model updates. Inference attacks are still possible, an extra layer of security was added to the system architecture. Privacy preserving methods were used to guarantee data privacy: Adaptive Differential Privacy and an efficient consensus protocol based on gradient verification. Differential privacy technology consists of adding a suitable noise to the updated model parameters to prohibit inference attacks. Gradient verification detects and filters malicious gradients injected into aggregation process. The model accuracy reaches 84%. It helps to reduce the privacy budget consumption and to resist to poisoning attacks. However, an unsuitable added noise may increase computing resources consumption to extract the original data.

[60] presented a FL method that uses Blockchain for a decentralized data management. The model purpose is to detect tuberculosis infection in a privacy preserving environment. A CNN model was trained on respectively 50, 70, 90, 110 X-ray images coming from MIoT devices. Blockchain was used to share model updates. The results showed that classifier effectiveness depends on the number of samples used for the training. The accuracy increased with an important number of samples and reached 75%. However, privacy preserving measures are still needed to be taken to ameliorate the classifier performance.

The work of [61] used a blockchain-based federated learning approach to remotely monitor patients' health using medical internet of things (MIOT). The model used a source data called Fitbit. The blockchain provided a secure sharing of patient health records. However, running a blockchain transaction network on a small device is challenging due to its high processing and bandwidth requirements.

Table 3 provides a summary of BCFL applications in medical field.

4. Proposed approach

To fulfill the previously announced objectives, the research's framework is divided into three main phases. As shown in Fig. 4 each phase's output is the input of the next phase.

4.1. Phase one: Dataset selection

Security and privacy are main objectives of this research. Medical field is one of the domains necessitating privacy preservation and high security. Therefore, Pima Indians Diabetes, a medical dataset, was used to implement the study. Pima dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. Based on diagnostic measurements, it is used to predict whether a

Table 3
Blockchain-based federated learning applications in medical field.

Ref	ML technique	Blockchain	Dataset	Results	Limitations
[53]	Deep learning algorithms for covid detection Biometric module: CNN for face recognition	ETHEREUM Hyperledger IPFS for off-chain	https://archive.ics.uci.edu/ml/datasets/spambase	Average accuracy: 91%	The privacy budget increases while the model accuracy increases. The DP guarantee decreases, while the number of iterations in the communication round in FL increases As the demand for energy increases, the score of IoHT sensors decreases.
[58]	ANN (2 hidden layers with 32 and 16 neurons respectively and a binary output layer)	Ethereum	https://www.kaggle.com/uciml/pima-indians-diabetes-database	Average accuracy: 73%	A model used in the medical field must have high accuracy. Low performances model could expose human's life to danger. The performances of the model need to be ameliorated.
[57]	Modified Capsule Network MLP: VGG16, AlexNet, Inception V3, ResNet 50-152 layers, MobileNet, DenseNet	Proof of work consensus algorithm	CC-19 dataset: 34.006 Computed Tomography (CT) scan slices for 89 persons (28.395 CT scan slices among them are of positive COVID-19 patients.)	Capsule Network: Accuracy 98.68% VGG16 Accuracy : 82.94%	
[60]	CNN	–	110 images of Tuberculosis Chest X-ray Image Datasets	Average accuracy: 75%	The privacy is preserved on behalf the model performances. The model need to be ameliorated to achieve more efficient results.
[59]	CNN	A small blockchain prototype based on Ethereum	https://www.kaggle.com/uciml/pima-indians-diabetes-database	Average Accuracy: 81.44%	The DP noise added on the gradient affects the model accuracy.
[61]	FedProx	A blockchain prototype based on Ethereum	https://www.kaggle.com/datasets/singhakash/fitbit-dataset		Running a blockchain transaction network on a small device is challenging due to its high processing and bandwidth requirements.

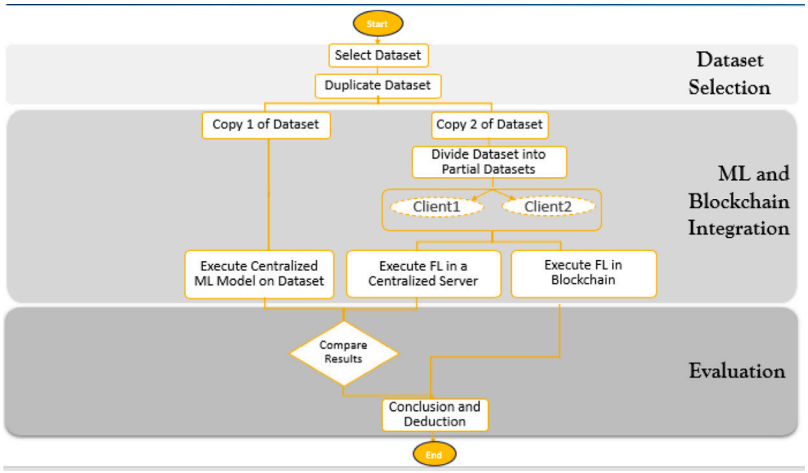

Fig. 4. Overview of overall framework.

patient is diabetic or not. It holds 768 labeled records of 21 or older females. It is composed of several medical predictor variables and one target variable, Outcome. Predictor variables consist of the number of pregnancies the patient has had, her Body Mass Index (BMI), Insulin level, age, and so on. A dataset summary description is provided in Table 4. The selected dataset is then replicated into two copies for the next framework execution phase.

4.2. Phase two: ML and blockchain integration

At this stage, three processes take place: centralized ML, FL and Blockchain integration.

4.2.1. Centralized ML

The dataset is a labeled dataset that traits a classification problem. Therefore, the ML model used in this study is a classifier model: a Multilayer Perceptron (MLP) model. MLP is a deep learning model. It is a feedforward Neural Network type [62]. It is composed of nodes organized into numerous layers to make prediction conclusions [63].

4.2.2. Federated learning

In a FL process, each client trains the model locally with its own data. The data is gathered by IoT sensors to monitor medical variables measurements such as Insulin level, blood pressure, Glucose level and so on. Based on these variables, the model predicts whether the person is diabetic or not. To simulate HIoT data, we divided the initial dataset into two partial datasets. Each dataset is associated to a client. The two datasets are, consequently, homogeneous and have the same features. The MLP training is, then, federated across two clients over the selected datasets and takes two rounds. Despite the fact that FL protects privacy, it could be subject to attacks at different levels: at the server side, between client and server, and at the client side. In order to secure the FL against attacks, Blockchain is used to protect it.

4.2.3. Blockchain integration

Instead of using a centralized server to manage the FL process, the study deploys Blockchain technology. A malicious server could alter model parameters and falsify the aggregation . Thus, Blockchain is

Table 4
Summary description of pima dataset features.

Feature Name	Description	Datatype
Pregnancies	Number of times pregnant	Integer
Glucose	Plasma glucose concentration a 2 h in an oral glucose tolerance test	Integer
blood Pressure	Diastolic blood pressure (mm Hg)	Integer
skin Thickness	Triceps skin fold thickness (mm)	Integer
Insulin	2-Hour serum insulin (mu U/ml)	Integer
BMI	Body mass index (weight in kg/(height in m^2)	Decimal
Diabetes Pedigree Function	Diabetes pedigree function	Decimal
Age	Age (years)	Integer
Outcome	Class variable (0 or 1)	Integer

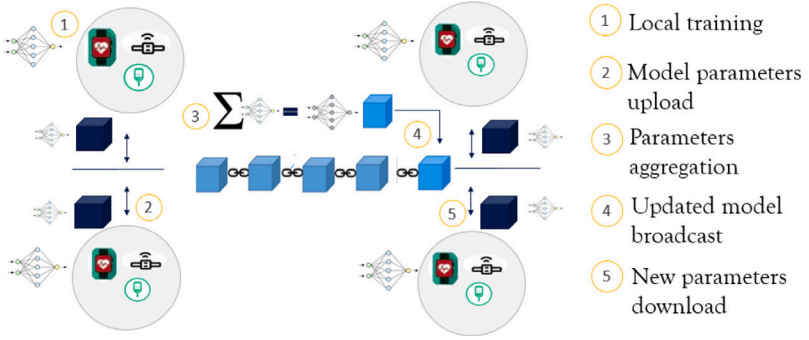

Fig. 5. Blockchain-based Federated Learning Implementation steps.

used to immunize model parameters against attacks. The aggregation process is implemented in a SC. The SC receives the parameters of the model from the distributed ledgers, aggregates them and sends them back to the clients' associated ledgers. The Blockchain-based Federated Learning implementation steps are presented in Fig. 5

4.3. Phase three: Evaluation

At the end of previous steps, the results of all experiments are collected and analyzed. Different performance measurements are used to evaluate the model. At this final phase of the research framework, the performances of the different processes are evaluated based on Accuracy, Precision and Recall.

- **TP** = **T**rue **P**ositive
- **TN** = **T**rue **N**egative
- **FP** = **F**alse **P**ositive
- **FN** = **F**alse **N**egative

Accuracy: It describes the ratio of correctly predicted results among all results. It tells how good the model was overall.

$$\textbf{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

Precision: It is the ratio of true positive results among total positive results. It tells how performance the model was in detecting a specific class .

$$\textbf{Precision} = \frac{TP}{TP + FP} \quad (2)$$

Recall: Called also Sensitivity, it describes the ratio of correctly predicted positive results among true positive results. Recall is most important in medical field. In fact, not only we need to treat the correctly positive results, but also cover the incorrectly predicted negative results. Therefore, the Recall must be of high value to classify the model as efficient.

$$\textbf{Recall} = \frac{TP}{TP + FN} \quad (3)$$

Table 5 shows the specifications of the development tools and the development computer system.

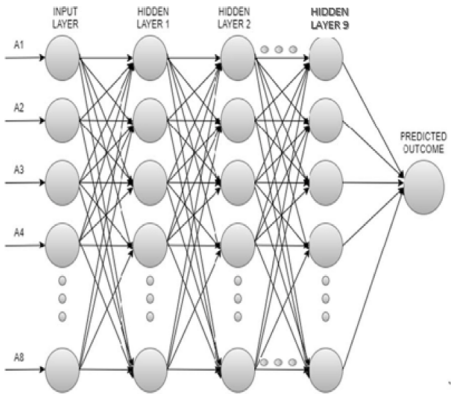

Fig. 6. MLP architecture: Input layer, hidden layers and output layer.

5. Experiment and results

This section is divided into two subsections. The first subsection describes the selected dataset and the ML model used in this research. The second subsection outlines the results achieved along this study.

5.1. Experiment

This subsection describes the selected dataset and the ML model used in this research. It also provides a description of the FL process and the Blockchain integration.

- **Dataset Selection and Preparation:**
  Pima Indians Diabetes dataset is used to implement the study; it is available at https://kaggle.com/datasets/uciml/pima-indians-diabetes-database
  The selected dataset is replicated into two copies: Copy1, Copy2. Copy1 is dedicated for centralized learning where Copy2 is dedicated for federated learning.
- **Centralized ML:**

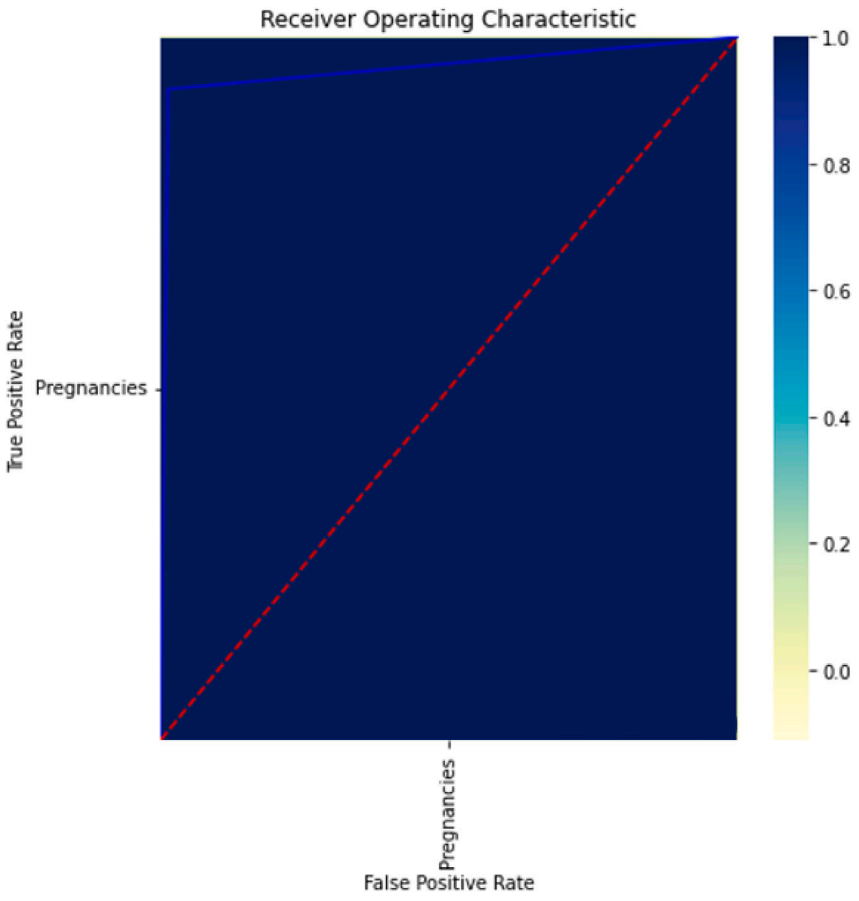

Fig. 7. ROC curve for MLP performance.

The MLP used in this research is composed of a 8-dimensions input layer, 9 hidden layers and a 1-dimension output layer. It relies on "adam" optimiser and on two activation functions: Rectifier Linear Unit (reLU) and Sigmoid. It uses K-fold cross-validation technique, the data is partitioned into 10 subsets. The model is trained on the first copy of the dataset (Copy1) for 20 times, each time using a different fold as the validation set and the remaining folds as the training set. Fig. 6 describes the used MLP model.

- **Federated Learning:**
  To simulate IoT datasets for FL, the second copy of the dataset Copy2 is divided into two datasets: Pima1, Pima2. The original dataset holds 768 records. If it gets divided into two equal datasets, the model would be trained on insufficient data. Therefore, to get effective results, the dataset was not equally divided because it is not voluminous. Pima 1 has a dataset size of 520, while Pima 2 has a dataset size of 550. The FL process involves two clients. It is conducted in two rounds. Each client trains the model on its own data. By the end of the round, it sends the model parameters back to the server for aggregation. It sends the weights and the biases of the model. The parameters are aggregated and sent back to clients. The model is set with the new parameters. Then, another round takes place until reaching the determined number of rounds.
- **Blockchain Integration:**
  Instead of using a centralized server to control the FL process, the study deploys Blockchain technology. SCs are used to store and aggregate the global and the local updates. The SC receives the parameters of the model from the distributed ledgers, aggregates them and sends them back to the clients' associated ledgers. Solidity programming language is used to develop SCs. Two main procedures are implemented: one for biases aggregation and the other for weights aggregation.

Table 5
Ecosystem Specifications.

Item	Specifications
Computer	OS: Windows10/ CPU: Intel Core i7- 2.60 GHz /RAM: 16 GB
Ganache	v2.5.4
Solidity	0.5.16
Metamask	10.18.0
Python	3.10

Table 6
MLP Performance.

Classifier	Dataset	Accuracy	Precision	Recall	Time (s)
MLP	Pima	97.11%	96.15%	92.59%	45.752

5.2. Results

This subsection presents the results of the research. First, it presents the results of the centralized ML execution on the first copy of the dataset. Then it provides the results of the FL execution on the partial datasets. These results are then compared to evaluate their performances . Finally, it presents the results of Blockchain integration.

- **Centralized ML Results:**
  The first copy of the dataset was used to train a Multi-Layer Perceptron in a centralized way. The performance of the model was measured using the standard measurements of accuracy, precision and recall. Table 6 summarizes the results of the model.

Table 7
First round MLP performances.

Classifier	Dataset	Accuracy	Precision	Recall	Time (s)
MLP	Pima1	94.23%	92.00%	85.18%	73.482
MLP	Pima2	95.45%	95.83%	85.18%	77.161

Table 8
Last round MLP performances.

Classifier	Dataset	Accuracy	Precision	Recall	Time (s)
MLP	Pima1	92.72%	88.09%	92.50%	98.235
MLP	Pima2	95.19%	92.68%	95.00%	107.475

Table 9
Parameters' aggregation cost in Gwei.

Procedure	Execution cost (gwei)
SC Deployment	4.324.313
Parameters Aggregation	5.698.860

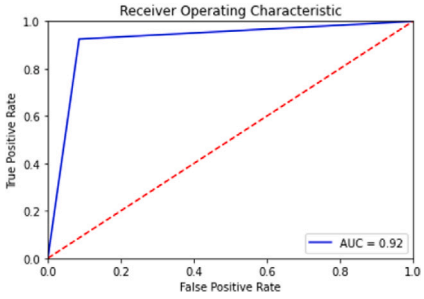

Fig. 8. Pima 1 ROC curve.

ROC (Receiver Operating Characteristic) curve is another way to measure the performance of the model. Fig. 7 depicts the ROC curve of the model.

- **Federated Learning Results:**
  Pima dataset was split into two datasets: Pima1, Pima2 . The FL process involved two clients. It took two rounds. The results of the model during the first round on each of the two datasets are represented in Table 7. By the end of the first round, the model is set with the new values. A new round took place and so on. The results of the last round are presented in Table 8. The ROC curves of the final round on both datasets are depicted in Figs. 8 and 9. Both ROC curves show satisfying performances. The model performs better on the dataset "Pima2" than "Pima1". This difference is due to the fact that "Pima2" is more voluminous than Pima1". Moreover, there is a class imbalance in the datasets; the model tends to focus more on the majority class to optimize the overall accuracy. Consequently, it achieves high recall for the positive class but compromises on correctly predicting the negative class, leading to low accuracy and precision.
- **Blockchain Integration Results:**
  The cryptocurrency used on Ethereum is Ether (ETH). Wei is the smallest denomination of Ether. To occur through Ethereum network, a transaction needs to be powered by some "gas". The gas price of a transaction costs Gwei. Gwei (a blend of giga and wei) is a denomination of ETH. One ETH is equal to one billion gwei. Table 9 shows the cost of the SC in gwei as computed by MetaMask during implementation.

6. Results discussion

The findings of the research are discussed in this section. It examines the differences between the outcomes of centralized ML and FL. These

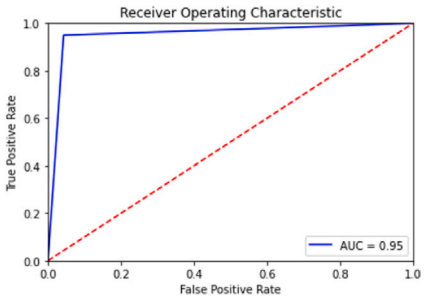

Fig. 9. Pima 2 ROC curve.

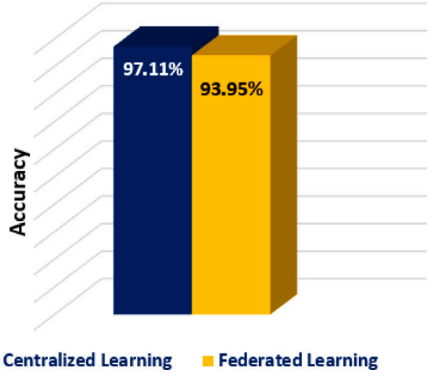

Fig. 10. The centralized MLP accuracy vs FL average accuracy.

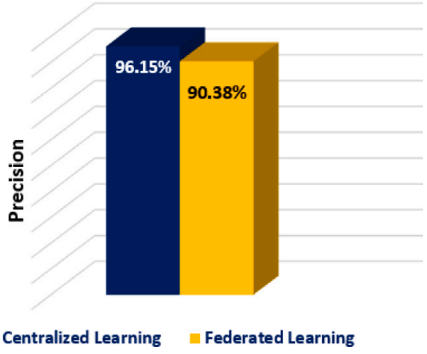

Fig. 11. The centralized MLP precision vs FL average precision.

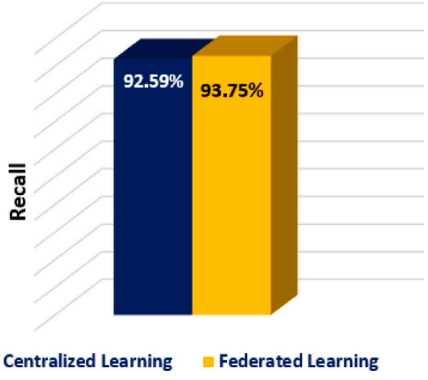

Fig. 12. MLP Centralized ML recall vs FL Average recall.

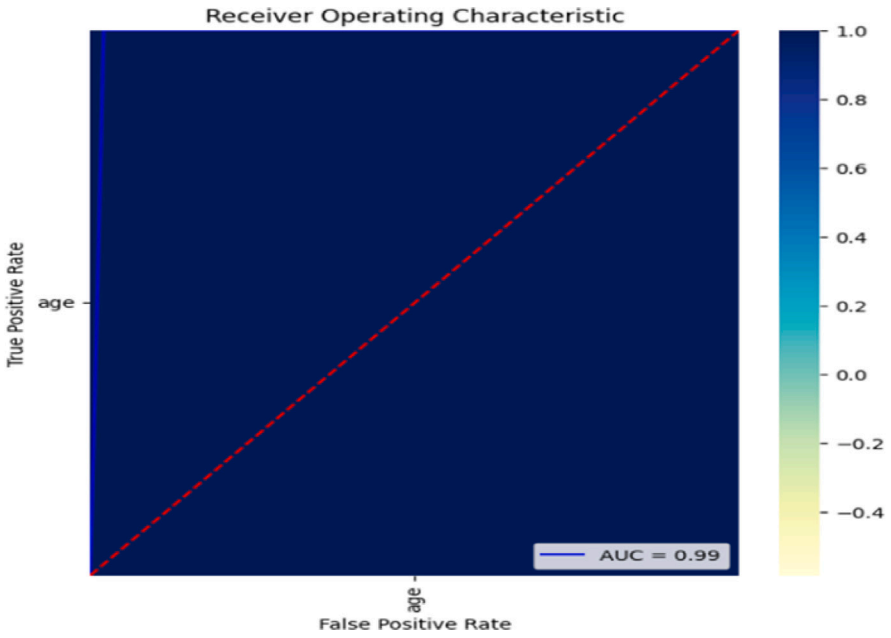

Fig. 13. ROC Curve for MLP performance on heart disease dataset.

findings are represented by three figures. The accuracy of the centralized ML and FL on both datasets is depicted in Fig. 10. The precision of the centralized ML and FL is depicted in Fig. 11. The recall of the centralized ML and FL can be seen in Fig. 12.

Comparing the results of centralized Learning and FL, it can be noticed that there is not a big difference between the two. When compared to the accuracy of the centralized ML, the average accuracy of the FL reduced by 3.16%. The slight decrease of model performances in FL is compensated by privacy preservation.

To evaluate the scalability of the proposed mechanism to large datasets, we applied it on Heart Disease Dataset, available at: https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset?resource=download&select=heart.csv

The used dataset contains 1025 records. It is composed of 14 features. The experiments show very encouraging results as the ROC curve in Fig. 13 shows. The model achieves a centralized accuracy of 99% and an average federated accuracy exceeding 95%. These results make the proposed approach transferable to various medical classification problems. It can be applied to additional classification tasks where accurate and efficient diagnosis and privacy preservation are a concern.

Clients are encouraged to employ a secured FL model that has good performance and a guarantee of privacy if they want their prediction requests to be fulfilled by that model. The protection of users' privacy in FL causes a moderate drop in model's performances but they are still efficient. In addition to that, the implementation of Blockchain technology brought an increased level of safety to the proposed system. It is essential that blockchain technology be integrated into intelligent healthcare systems in order to boost patients' trust. In the sphere of medicine, the application of such technologies would significantly improve the linkage of Blockchain and FL. Additional intelligent healthcare applications may stand to gain from Blockchain's potentially significant contribution and see their services improved as a result. Therefore, Clients would use the applications provided by The HIoT without worrying about their privacy or the effectiveness of the services provided. That would be a big help in terms of saving more human lives and providing new prospects for a better future in terms of e-health.

Our proposed solution is reliable enough to allow the sharing of private medical data and the exploitation of opportunities presented by smart healthcare. Applications for the IoT built on top of such

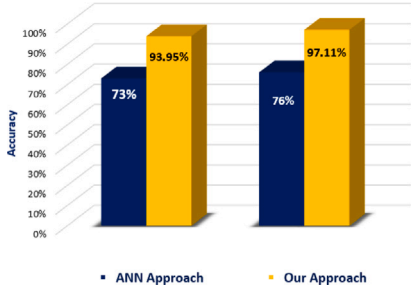

Fig. 14. The figure compares the performances of the ANN approach and our approach performances.

intelligent systems would significantly improve the quality of life for those living in those areas. As a consequence of this, the Blockchain-based federated learning system appears to be an ambitious solution for the problem that has been investigated, given that it preserves users' privacy and guarantees their safety. The work of [58] developed and built a Blockchain-based federated learning system to predict diabetes. It trained an ANN on the same dataset . The FL average accuracy that it attained was 73%, which contrasts with the FL average accuracy that our approach achieved, which was 93.95 (see Fig. 14).

7. Threats to validity

Some challenges and obstacles were encountered during the conduction of this study. They can be outlined as follows:

1. The technology of Blockchain and the FL are relatively new. Applying these technologies in the medical field is still limited. That is why finding sufficient literature was a challenge.
2. The dataset used in this study is not an IoT dataset. To simulate IoT datasets, we needed to divide the initial dataset into smaller datasets.
3. The functionalities of SC programming languages are limited.

8. Conclusion

Some noteworthy achievements have been endorsed by this study. In this paper, we systematically analyzed previous research on the

subject of using FL and Blockchain in healthcare. The study revealed that there is an increasing need for privacy and security measures in the healthcare sector. The second goal of the study was leveraging MLP classifier for predicting diabetes. It is essential that a medical model have excellent performance. The proposed method outperforms current state-of-the-art methods in terms of accuracy and privacy preservation, according to experimental results on a public diabetes dataset. The accuracy, precision, and recall of the deployed model are all outstanding. The implemented model has a satisfying accuracy that is near the centralized model's accuracy. In addition to that, by using Blockchain, the aggregation process implemented in a SC was secured against attacks and falsification. Future research could include more ML algorithms and unsupervised learning techniques. In addition, the addition of a sizable number of IoT datasets has the potential to significantly boost the model's accuracy. A more robust model can be built with more and different types of data. Further, the FL process may use a security approach that consists of promoting safe clients and rejecting harmful clients from the training process in order to improve security. Further, while it is important to point out Blockchain's role in keeping the system secure, the latter may be vulnerable to attacks. Moreover, we tend to work on reducing the Blockchain high latency.

Declaration of competing interest

We have no conflicts of interest to disclose.

Data availability

The data selected for the experiment is freely available at https://kaggle.com/datasets/uciml/pima-indians-diabetes-database.

References

[1] P. Mikalef, K. Conboy, J.E. Lundström, A. Popovič, Thinking responsibly about responsible AI and 'the dark side'of AI, Eur. J. Inf. Syst. 31 (3) (2022) 257–268.

[2] G. Hcini, I. Jdey, H. Ltifi, Improving malaria detection using L1 regularization neural network, JUCS: J. Univ. Comput. Sci. 285 (1087) (2022) 10–1107.

[3] B. Hrnjica, D. Music, S. Softic, Model-based recommender systems, in: Trends Cloud-Based IoT, Springer, 2020, pp. 125–146.

[4] C. Badue, R. Guidolini, R.V. Carneiro, P. Azevedo, V.B. Cardoso, A. Forechi, A.F. De Souza, Self-driving cars: A survey, Expert Syst. Appl. 165 (2021) 113816.

[5] W. Li, Y. Chai, F. Khan, S.R.U. Jan, S. Verma, V.G. Menon, X. Li, et al., A comprehensive survey on machine learning-based big data analytics for IoT-enabled smart healthcare system, Mob. Netw. Appl. 26 (1) (2021) 234–252.

[6] B. Rouet-Leduc, C. Hulbert, N. Lubbers, K. Barros, C.J. Humphreys, P.A. Johnson, Machine learning predicts laboratory earthquakes, Geophys. Res. Lett. 44 (18) (2017) 9276–9282.

[7] I. Jdey, et al., Fuzzy fusion system for radar target recognition, Int. J. Comput. Appl. Inf. Technol. 1 (3) (2012) 136–142.

[8] I. Jdey, G. Hcini, H. Ltifi, Deep learning and machine learning for malaria detection: Overview, challenges and future directions, Int. J. Inf. Technol. Decis. Mak. (2022).

[9] L. Awassa, I. Jdey, H. Dhahri, G. Hcini, A. Mahmood, E. Othman, M. Haneef, Study of different deep learning methods for coronavirus (COVID-19) pandemic: Taxonomy, survey and insights, Sensors 22 (5) (2022) 1890.

[10] H. Habibzadeh, K. Dinesh, O.R. Shishvan, A. Boggio-Dandry, G. Sharma, T. Soyata, A survey of healthcare Internet of Things (HIoT): A clinical perspective, IEEE Internet Things J. 7 (1) (2019) 53–71.

[11] P. Samangouei, M. Kabkab, R. Chellappa, Defense-GAN: Protecting classifiers against adversarial attacks using generative models, 2018, arXiv preprint arXiv:1805.06605.

[12] I. Jdey, Trusted smart irrigation system based on fuzzy IoT and blockchain, in: International Conference on Service-Oriented Computing, Springer Nature Switzerland, Cham, 2022, pp. 154–165.

[13] Z. El Ouazzani, H. El Bakkali, S. Sadki, Privacy preserving in digital health: Main issues, technologies, and solutions, in: Research Anthology on Privatizing and Securing Data, IGI Global, 2021, pp. 1503–1526.

[14] S. Tian, W. Yang, J.M. Le Grange, P. Wang, W. Huang, Z. Ye, Smart healthcare: Making medical care more intelligent, Global Health J. 3 (3) (2019) 62–65.

[15] G. Eysenbach, et al., What is e-Health? J. Med. Internet Res. 3 (2) (2001) e833.

[16] E.R. Dorsey, E.J. Topol, State of telehealth, New Engl. J. Med. 375 (2) (2016) 154–161.

[17] J. Craig, V. Petterson, Introduction to the practice of telemedicine, J. Telemed. Telecare 11 (1) (2005) 3–9.

[18] M.A. Akkaş, R. Sokullu, H.E. Cetin, Healthcare and patient monitoring using IoT, Internet Things 11 (2020) 100173.

[19] K.K. Patel, S.M. Patel, P. Scholar, Internet of things-IoT: Definition, characteristics, architecture, enabling technologies, application & future challenges, Int. J. Eng. Sci. Comput. 6 (5) (2016).

[20] M. Maksimović, V. Vujović, Internet of Things based e-Health systems: Ideas, expectations and concerns, in: Handbook of Large-Scale Distributed Computing in Smart Healthcare, Springer, 2017, pp. 241–280.

[21] I. Zagan, V.G. Gaitan, A.-I. Petrariu, A. Brezulianu, Healthcare IoT m-GreenCARDIO remote cardiac monitoring system–concept, theory of operation and implementation, Adv. Electr. Comput. Eng. 17 (2) (2017) 23–30.

[22] F. Lamonaca, E. Balestrieri, I. Tudosa, F. Picariello, D.L. Carnì, C. Scuro, F. Bonavolontà, V. Spagnuolo, G. Grimaldi, A. Colaprico, An overview on Internet of Medical Things in blood pressure monitoring, in: 2019 IEEE International Symposium on Medical Measurements and Applications, MeMeA, IEEE, 2019, pp. 1–6.

[23] S.D. Machado, J.E.d.R. Tavares, M.G. Martins, J.L.V. Barbosa, G.V. González, V.R.Q. Leithardt, Ambient intelligence based on iot for assisting people with Alzheimer's disease through context histories, Electronics 10 (11) (2021) 1260.

[24] Z. Alansari, S. Soomro, M.R. Belgaum, S. Shamshirband, The rise of Internet of Things (IoT) in big healthcare data: review and open research issues, in: Progress in Advanced Computing and Intelligent Engineering, Springer, 2018, pp. 675–685.

[25] T.N. Gia, I. Tcarenko, V.K. Sarker, A.M. Rahmani, T. Westerlund, P. Liljeberg, H. Tenhunen, IoT-based fall detection system with energy efficient sensor nodes, in: 2016 IEEE Nordic Circuits and Systems Conference, NORCAS, IEEE, 2016, pp. 1–6.

[26] L. Liu, J. Xu, Y. Huan, Z. Zou, S.-C. Yeh, L.-R. Zheng, A smart dental health-IoT platform based on intelligent hardware, deep learning, and mobile terminal, IEEE J. Biomed. Health Inf. 24 (3) (2019) 898–906.

[27] M.U. Farooq, M. Waseem, S. Mazhar, A. Khairi, T. Kamal, A review on Internet of Things (IoT), Int. J. Comput. Appl. 113 (1) (2015) 1–7.

[28] A.M. Mzahm, M.S. Ahmad, A.Y.C. Tang, Agents of things (AoT): An intelligent operational concept of the Internet of Things (IoT), in: 2013 13th International Conference on Intellient Systems Design and Applications, IEEE, Salangor, Malaysia, 2013, pp. 159–164.

[29] John-P.-Mello-Jr., 11 top cloud security threats, 2022, https://www.csoonline.com/article/3043030/top-cloud-security-threats.html?page=2.

[30] V. Mothukuri, R.M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, G. Srivastava, A survey on security and privacy of federated learning, Future Gener. Comput. Syst. 115 (2021) 619–640.

[31] X. Yin, Y. Zhu, J. Hu, A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions, ACM Comput. Surv. 54 (6) (2021) 1–36.

[32] T. Moulahi, R. Jabbar, A. Alabdulatif, S. Abbas, S. El Khediri, S. Zidi, M. Rizwan, Privacy-preserving federated learning cyber-threat detection for intelligent transport systems with blockchain-based security, Expert Syst. (2022) e13103.

[33] S. AbdulRahman, H. Tout, H. Ould-Slimane, A. Mourad, C. Talhi, M. Guizani, A survey on federated learning: The journey from centralized to distributed on-site learning and beyond, IEEE Internet Things J. 8 (7) (2020) 5476–5497.

[34] S.A. Rahman, H. Tout, C. Talhi, A. Mourad, Internet of Things intrusion detection: Centralized, on-device, or federated learning? IEEE Netw. 34 (6) (2020) 310–317.

[35] D.C. Nguyen, M. Ding, Q.V. Pham, P.N. Pathirana, L.B. Le, A. Seneviratne, H.V. Poor, Federated learning meets blockchain in edge computing: Opportunities and challenges, IEEE Internet Things J. 8 (1280) (2021) 166–12825.

[36] Z. Wang, Q. Hu, Blockchain-based federated learning: A comprehensive survey, 2021, arXiv preprint arXiv:2110.02182.

[37] X. Luo, Y. Wu, X. Xiao, B.C. Ooi, Feature inference attack on model predictions in vertical federated learning, in: 2021 IEEE 37th International Conference on Data Engineering, ICDE, IEEE, 2021, pp. 181–192.

[38] L. Lyu, H. Yu, Q. Yang, Threats to federated learning: A survey, 2020, arXiv preprint arXiv:2003.02133.

[39] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, Y. Gao, A survey on federated learning, Knowl.-Based Syst. 216 (2021) 106775.

[40] A.B. Levina, V.Y. Kadykov, D.I. Kaplun, New direction in cryptography: Homomorphic encryption, in: 2021 International Conference Automatics and Informatics, ICAI, IEEE, 2021, pp. 234–237.

[41] S. Zhou, M. Liao, B. Qiao, X. Yang, A survey of security aggregation, in: 2022 24th International Conference on Advanced Communication Technology, ICACT, IEEE, 2022, pp. 334–340.

[42] C. Ma, J. Li, L. Shi, M. Ding, T. Wang, Z. Han, H.V. Poor, When federated learning meets blockchain: A new distributed learning paradigm, IEEE Comput. Intell. Mag. 17 (3) (2022) 26–33.

[43] Y. Lu, The blockchain: State-of-the-art and research challenges, J. Ind. Inf. Integr. 15 (2019) 80–90.

[44] A. Narayanan, J. Bonneau, E. Felten, A. Miller, S. Goldfeder, Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction, Princeton University Press, 2016.
[45] D. Puthal, N. Malik, S.P. Mohanty, E. Kougianos, G. Das, Everything you wanted to know about the blockchain: Its promise, components, processes, and problems, IEEE Consum. Electron. Mag. 7 (4) (2018) 6–14.
[46] U. Mukhopadhyay, A. Skjellum, O. Hambolu, J. Oakley, L. Yu, R. Brooks, A brief survey of cryptocurrency systems, in: 2016 14th Annual Conference on Privacy, Security and Trust, PST, IEEE, 2016, pp. 745–752.
[47] D. Yaga, P. Mell, N. Roby, K. Scarfone, Blockchain technology overview, 2019, arXiv preprint arXiv:1906.11078.
[48] M. Di Pierro, What is the blockchain? Comput. Sci. Eng. 19 (5) (2017) 92–95.
[49] G. Srivastava, S. Dhar, A.D. Dwivedi, J. Crichigno, Blockchain education, in: 2019 IEEE Canadian Conference of Electrical and Computer Engineering, CCECE, IEEE, 2019, pp. 1–5.
[50] S. Namasudra, G.C. Deka, P. Johri, M. Hosseinpour, A.H. Gandomi, The revolution of blockchain: State-of-the-art and research challenges, Arch. Comput. Methods Eng. 28 (3) (2021) 1497–1515.
[51] Z. Zheng, S. Xie, H.-N. Dai, W. Chen, X. Chen, J. Weng, M. Imran, An overview on smart contracts: Challenges, advances and platforms, Future Gener. Comput. Syst. 105 (2020) 475–491.
[52] C. Cachin, et al., Architecture of the hyperledger blockchain fabric, in: Workshop on Distributed Cryptocurrencies and Consensus Ledgers, Vol. 310, no. 4, Chicago, IL, 2016, pp. 1–4.
[53] M.A. Rahman, M.S. Hossain, M.S. Islam, N.A. Alrajeh, G. Muhammad, Secure and provenance enhanced internet of health things framework: A blockchain managed federated learning approach, Ieee Access 8 (2020) 205071–205087.
[54] S. Jlassi, I. Jdey, H. Ltifi, Bayesian hyperparameter optimization of deep neural network algorithms based on ant colony optimization, in: Document Analysis and Recognition–ICDAR 2021: 16th International Conference, Lausanne, Switzerland, September (2021) 5–10, Proceedings, Part III 16, Springer International Publishing, 2021.
[55] N. Slimani, I. Jdey, M. Kherallah, Performance comparison of machine learning methods based on CNN for satellite imagery classification, in: 2023 9th International Conference on Control, Decision and Information Technologies, CoDIT, Rome, Italy, 2023, pp. 185–189.
[56] I. Jdey, et al., The contribution of fusion techniques in the recognition systems of radar targets, 2012, p. 94.
[57] R. Kumar, A.A. Khan, J. Kumar, N.A. Golilarz, S. Zhang, Y. Ting, C. Zheng, W. Wang, et al., Blockchain-federated-learning and deep learning models for covid-19 detection using CT imaging, IEEE Sens. J. 21 (14) (2021) 16301–16314.
[58] O. El Rifai, M. Biotteau, X. de Boissezon, I. Megdiche, F. Ravat, O. Teste, Blockchain-based federated learning in medicine, in: International Conference on Artificial Intelligence in Medicine, Springer, 2020, pp. 214–224.
[59] Y. Chang, C. Fang, W. Sun, A blockchain-based federated learning method for smart healthcare, Comput. Intell. Neurosci. 2021 (2021).
[60] D. Połap, G. Srivastava, A. Jolfaei, R.M. Parizi, Blockchain technology and neural networks for the Internet of Medical Things, in: IEEE INFOCOM 2020-IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPS, IEEE, 2020, pp. 508–513.
[61] K. Farooq, H.J. Syed, S.O. Alqahtani, W. Nagmeldin, A.O. Ibrahim, A. Gani, Blockchain federated learning for in-home health monitoring, Electronics 12 (1) (2022) 136.
[62] H. Taud, J. Mas, Multilayer perceptron (MLP), in: Geomatic Approaches for Modeling Land Change Scenarios, Springer, 2018, pp. 451–455.
[63] G. Hcini, I.M.E.N. Jdey, A. Heni, H. Ltifi, Hyperparameter optimization in customized convolutional neural network for blood cells classification, J. Theor. Appl. Inf. Technol 99 (2021) 5425–5435.