applied sciences

MDPI

Article

Hybrid Classifier-Based Federated Learning in Health Service Providers for Cardiovascular Disease Prediction

Muhammad Mateen Yaqoob [1,*], Muhammad Nazir [1], Muhammad Amir Khan [2,*], Sajida Qureshi [3] and Amal Al-Rasheed [4]

[1] Department of Computer Science, HITEC University Taxila, Taxila 47080, Pakistan
[2] Department of Computer Science, COMSATS University Islamabad Abbottabad Campus, Abbottabad 22060, Pakistan
[3] Computer Science Department, Faculty of Computer Sciences, ILMA University, Karachi 75190, Pakistan
[4] Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
* Correspondence: mateenyaqoob@gmail.com (M.M.Y.); amirkhan@cuiatd.edu.pk (M.A.K.)

Abstract: One of the deadliest diseases, heart disease, claims millions of lives every year worldwide. The biomedical data collected by health service providers (HSPs) contain private information about the patient and are subject to general privacy concerns, and the sharing of the data is restricted under global privacy laws. Furthermore, the sharing and collection of biomedical data have a significant network communication cost and lead to delayed heart disease prediction. To address the training latency, communication cost, and single point of failure, we propose a hybrid framework at the client end of HSP consisting of modified artificial bee colony optimization with support vector machine (MABC-SVM) for optimal feature selection and classification of heart disease. For the HSP server, we proposed federated matched averaging to overcome privacy issues in this paper. We tested and evaluated our proposed technique and compared it with the standard federated learning techniques on the combined cardiovascular disease dataset. Our experimental results show that the proposed hybrid technique improves the prediction accuracy by 1.5%, achieves 1.6% lesser classification error, and utilizes 17.7% lesser rounds to reach the maximum accuracy.

Keywords: heart disease prediction; hybrid technique; ABC-SVM; privacy-aware machine learning; intelligence-based healthcare

check for updates

Citation: Yaqoob, M.M.; Nazir, M.; Khan, M.A.; Qureshi, S.; Al-Rasheed, A. Hybrid Classifier-Based Federated Learning in Health Service Providers for Cardiovascular Disease Prediction. *Appl. Sci.* **2023**, *13*, 1911. https://doi.org/10.3390/app13031911

Academic Editors: Lucian Mihai Itu, Constantin Suciu and Anamaria Vizitiu

Received: 2 January 2023
Revised: 29 January 2023
Accepted: 30 January 2023
Published: 1 February 2023

1. Introduction

The Internet of Things (IoT) enables the connectivity of physical objects and computational power so they may connect to the Internet. The IoT has the potential to assist in the development of applications that are both adaptable and efficient across a variety of industries, including healthcare, environmental monitoring, and industrial control systems. The IoT in the healthcare environment has led to the establishment of the Internet of Medical Things (IoMT), a cutting-edge area of a cyber physical system for wellness and wellbeing. Integrating these solutions into the HSP system has the potential to improve care services, quality of life, and open the door to cost-effective solutions [1]. For further analysis, the biomedical information pertaining to people is obtained. This information includes medical records, photographs, physiological signals, and many more forms. Given that the IoMT's cyber physical system collects data from several users, the volumetric scale of this biomedical data is enormous [2]. Smartwatches, wristbands, and other wearable sensing devices, among others, help in early illness diagnosis and warning. These wearable devices include strong and application-specific computational architecture that is housed in a distant HSP cloud data center, enhancing their capabilities (for real-time and early detection of health concerns). In IoMT-based healthcare solutions, wearable devices are often collected at the HSP's data center with the goal of preventing, diagnosing, and treating a variety of human health-related issues including cardiovascular illnesses (CVD). The construction of an effective electronic healthcare infrastructure is difficult due to the vast number of data that are gathered from multiple sources, including end users and other stakeholders in the delivery of health services.

According to World Health Organization (WHO) projections, CVD-related mortality accounts for close to 18 million fatalities annually globally [3]. Numerous risk factors, such as a history of heart attack, obesity, stress, high blood pressure, smoking, excessive use of alcohol, and high cholesterol, can all contribute to CVD. CVD impairs heart function and results in issues including strokes and impaired blood vessel function. The ability to treat CVD effectively and quickly is crucial for patient survival. The academic and industry communities are paying close attention to machine learning (ML)-based approaches for the accurate detection and prognosis of cardiac disorders. In [4], for example, a hybrid strategy based on a random forest and linear model was proposed to increase heart disease prediction accuracy. Authors in [5] proposed a feature selection and classification technique for identifying cardiac illness in an e-healthcare system.

To create an effective prediction model for tracking a patient's health state, traditional ML models are trained on vast amounts of user data. Although it is organized by several autonomous HSPs, these healthcare data are available in scattered, isolated silos. Even if there are a lot of aggregate data in different businesses, sharing the data is limited because of worries about security and privacy. Similarly, collected user data from the crowd are too restricted. These restrictions are enforced through regulatory laws such as the European Union's GDPR [6], China Cyber Security Law [7], and the United States' CCPA of 2018 [8]. Hence, it is not trivial to accumulate large amounts of user data in real-time healthcare applications to train powerful predictive models with high-quality training data. On the other hand, if the collection of user data is allowed, it is still not trivial to process these crowd-generated data, since the volume and velocity of the incoming data at the central server of HSP put a lot of burden on the network backhaul, delimited by the processing and storage capabilities of the central server. Indeed, with these restrictions in place, the number of training samples would not be large enough to generalize the model, affecting the performance of the trained model. To overcome these challenges, Google in [9,10] proposed federated learning (FL): a combination of distributed and incremental machine learning. FL is a distributed privacy-preserving machine learning technique that enables the collaborative training of a shared global predictive model without the need of uploading private local data to a central server to overcome the privacy concerns caused by centralized machine learning.

The FL algorithm's efficacy can be further improved by introducing feature selection at the distributed nodes. Feature selection will improve the identification of common features set in the sensory health data and distributed over the healthcare registries. Furthermore, feature selection will also help in dimensionality reduction to lower the computational cost and the model size. In this regard, recently, a feature-optimized federated learning-based technique was proposed in [11]; they addressed the issue of dimensionality reduction and communication efficiency for heart disease by improving the distributed nodes' learning technique. For the security and privacy issues in the cloud computing environment, federated learning incentive-based mechanisms were introduced in [12–14]. Recently, some meta-heuristic techniques have been proposed to further expand the solution search space for cloud-based healthcare systems [15,16]. These techniques also aim to minimize the fitness (objective) function though preserve the size of the population, increase weight adaptation rates, improve local search techniques, offer fitness function-improved computation, provide solutions to avoid local minima, and enhance the convergence rate of the algorithm.

Deep learning (DL) and SVM are both effective methods, although they are made to address distinct challenges. While DL is better suited for big datasets with many features, SVM works well for small to medium-sized datasets with few features. In comparison to

DL, SVM is less prone to overfitting and is known to be successful in high-dimensional space. SVM can be trained using various kernel functions, such as linear, polynomial, and radial basis functions (RBF), which can help to handle non-linear data. In tabular datasets, SVM can be used in conjunction with feature scaling, normalization, and dimensionality reduction techniques to improve the performance. Motivated by these highlighted issues of data privacy, improved feature selection, and classification for heart disease, in this manuscript, we proposed a privacy-aware FL-based framework that utilizes federated matched averaging at the HSPs' cloud end with a hybrid technique of modified artificial bee colony with support vector machine (MABC-SVM) for optimization for effective CVD prediction, respectively, at the client nodes. The M-ABC at the HSP client for optimal solution search works in four phases, i.e., the initialization phase, employed bee phase, onlooker bee phase, and the scout bee phase. These steps are described in detail in Section 4.1. The primary contributions of the proposal are enumerated as follows:

- An FL-based framework is proposed in this paper to overcome the problem of data privacy for HSP systems.
- We utilize the modified version of a federated matched averaging (FedMA) algorithm to preserve the privacy of heart disease data and to address the issues of the HSP's central model updation and communication efficiency.
- A hybrid technique comprised of a modified artificial bee colony and support vector machine (MABC-SVM) is proposed for the prediction of CVD with improved prediction accuracy. This hybrid algorithm is introduced at the client end of HSP.
- Our hybrid method's performance in terms of communication efficiency, classification error, and prediction accuracy is assessed and compared to current FL approaches.

The rest of the paper is organized as follows. Section 2 provides context for federated learning, the MABC method for optimum feature selection, and the SVM classification algorithm. Section 3 is an overview of relevant work. The proposed hybrid FL-based method is explained in Section 4. Section 5 is concerned with performance and the outcomes' evaluation. The conclusion and future work are presented in Section 6.

2. Background

In this section, we provide a brief overview of the methodologies that were used to build the FedMA with MABC-RB-SVM framework for privacy-aware heart disease prediction.

2.1. Basics of Federated Learning

To train a model, one needs access to data, which is the core of the area of artificial intelligence, and it frequently occurs in isolated data islands. The problem of isolated data silos is easily resolved by centralizing data processing. As international privacy protection laws for users strengthen, data collection for training models becomes more challenging. The issue of how to legally address data islands has sparked considerable discussion and research in the field of artificial intelligence. Traditional data analytics approaches are already at capacity due to the many rules that must be adhered to while attempting to address the data silo problem.

By jointly training algorithms without transferring the data, federated learning is a learning paradigm that aims to overcome the issues of data governance and privacy. While data are stored locally, federated learning trains statistical models across data silos. By retaining the data on the device, FL aims to provide a collaborative learning process that is privacy conscious and uses a shared model. As a result, users of FL will benefit from individualized machine learning that also addresses privacy concerns.

2.2. M-ABC-Based Optimization Algorithm

The modified artificial bee colony (M-ABC), a swarm intelligence-based technique, was proposed in [17]. The M-ABC is the upgraded version of artificial bee colony optimization, which is a method that mimics the intelligent foraging behavior of honeybee colonies to find the best solution to a problem. It simulates the bees' actions while they search for nectar, and keeps a group of potential answers, referred to as "bees", that explore the possible solutions and adjust their positions based on the quality of their findings. This method is particularly effective for solving complex optimization issues that are hard to resolve with traditional optimization methods. This M-ABC algorithm has three types of artificial bees such as onlooker, employed, and scout bee. While onlooker bees pick a source constructed on the employed bee's dance, scout bees oversee discovering new food sources. As a result of being connected to their food supply, the employed bees are shielded against exploitation. The observer bees and the scout bees are not connected to any food source. The primary goals of the fitness function are the best possible classification error and communication effectiveness of the models that are obtained from HSP sites. In order to increase accuracy, the fitness function seeks to reduce classification errors and round consumption.

In the M-ABC, the scout bee is combined with the firefly algorithm, and the modified technique is a combination of two different metaheuristic optimization techniques, namely, the artificial bee colony (ABC) and the firefly algorithm. The ABC algorithm generates a population of potential solutions, and the firefly algorithm improves the solutions by simulating the flashing behavior of fireflies. This hybrid algorithm is known for its global optimization abilities and has been applied to various optimization problems, showing better results than the original ABC or firefly algorithm alone. It is useful for solving complex optimization problems that are difficult to solve with traditional techniques.

2.3. SVM Classification Technique

The support vector machine (SVM) is widely used in intelligence-based systems for classification problems. The fundamental principle of the SVM classification algorithm is to discover the negative samples and the optimal selection for dividing positive samples. To attain the best generalization ability while remaining resistant to overfitting, the SVM attempts to determine the trade-off between lowering the training set error and maximizing the margin [18]. In addition, one of the best things about the SVM is that it uses convex quadratic programming, which gives only global minima and keeps the program from getting stuck in local minima.

The data are converted using a method called kernel trick using the SVM. The SVM kernel is a function that converts non-separable problems into separable problems by taking low-dimensional input space and transforming it into higher-dimensional space. Data conversion is used to determine the best splitting line among the expected outcomes. The border can range from a straightforward narrow margin for binary classes to a more challenging splitting including multiple classes [19].

3. Related Work

The contemporary machine learning research field is faced with two major challenges: data isolation and privacy and security issues. In methods utilizing standard ML, centralized training data are necessary. Around the world, laws are implemented to protect the privacy of data [6–8]. Therefore, the main difficulty for conventional machine learning algorithms is data privacy. A federated stochastic gradient descent (FedSGD) and federated averaging (FedAvg)-based technique first developed by Google in [9] offered some hope for overcoming these difficulties. In [10], a method based on FL was suggested to address the problems of data silos and privacy. To address the problems with data security in the conventional artificial intelligence field, they created an extensive architecture based on federated learning. Their suggested solution was divided into two categories: horizontal and vertical FL.

A description of the different machine learning deployment models, including centralized, distributed, and federated learning, was given by the authors in [20]. With careful consideration, they have described how machine learning architectures have developed. The authors of the research in [21] created a federated learning-based model for individuals with diseases that are likely to require hospitalization. They made use of data from electronic health records (EHRs) spread across various sources or agents. To use FL to solve the issue of large-scale sparse computing, the authors presented the clustering-based approach for dual splitting. Their suggested method yielded similar classifier prediction accuracy. The MNIST dataset was utilized by the authors of [22] to test and assess the three FL-based methods. A Bayesian correlated *t*-test was also employed. When client uploads were restricted to 10,000, FedAvg surpassed CO-OP and FSVRG algorithms, in their assessment. They have employed balanced data distribution, where each customer receives the same volume of information. The authors of [23] suggested a modified version of the standard FL with the aim of improving the algorithm's accuracy and convergence rate. To implement Bayesian non-parameterized approaches for heterogeneous data, they introduced the FedMA algorithm, which is a layer-wise version of the FL algorithm. Their suggested FedMA outperformed in terms of convergence, accuracy, and communication size reduction. The authors of [24] examined technical issues and other factors regarding the data privacy in the distributed implementation environment for FL algorithms. In their study, they outlined the features and results of a few of the optimization strategies for FL implementation. Additionally, they have discussed certain commercial consequences for federated learning that will be expected.

The authors of [25] have suggested an algorithm that distributes weights according to how much each class contributes to the local models. Using patients' chest x-ray data, machine learning-based algorithms can contribute to the identification of COVID-19. In contrast to conventional machine learning techniques, an FL version was suggested in [26] to discover COVID-19 with improved prediction accuracy. A blockchain-based approach based on federated learning was suggested by authors in [27] to address the problem of data privacy for IoMT-based healthcare systems. Their suggested solution was a hybrid strategy built on federated learning and the maximum approximation of the Gaussian mixture model, and it used blockchain to address the issue of user data privacy. Their suggested approach demonstrates that IoMT data training may be carried out utilizing local privacy to stop data leaking.

In the past, researchers gathered sensed data from HSP devices and then utilized that data to predict about several diseases. The authors of [28] suggested a version of FL with a Bayesian inference model to construct a privacy-aware heart rate prediction approach. Comparing this FedARX approach to conventional machine learning models, it achieves accurate and reliable heart rate prediction. A meta-heuristic method called artificial bee colony (ABC) optimization was suggested by the authors in [29] as a way to efficiently manage and optimize the calculation of offloading for IoT-based applications. Their method effectively controls the computing workload for IoT applications with limited resources. The authors in [30] suggested a fast-convergent technique that accomplishes intelligent selection of each device at every round of training the model in order to maximize the convergence speed of federated learning. To increase the convergence rate, their approach employs precise and efficient approximation for the transmission of a nearly optimum distribution of device selection.

Other approaches, such as a hybrid technique combining a linear discriminant analysis with modified ant lion optimization for classification [31], a gradient boosting decision tree with fuzzy logic algorithm [32], a hybrid of modified scalp swarm optimization and adaptive neuro-fuzzy inference system [33], and a multi-objective function using meta-heuristics [34], were also presented in the literature as strategies for predicting heart disease. The use of a hybrid classifier and a modified neural network with a deep learning focus was presented in [35–38] as a method for monitoring and predicting cardiac problems. Methods such as the Boltzmann-based model for higher order [39], modified hybrid method using classifiers and optimizers [40], two-stage-based localization of the classifiers [41], and hybrid classifier based on nave Bayes and random forest [42] were also anticipated to improve classification accuracy with less error.

4. Proposed Hybrid FL-Based Framework

This section suggests a federated learning architecture that addresses privacy concerns and effectively predicts cardiac disease in a healthcare system that is sensitive to privacy concerns. Table 1 provides a description of the symbols used in our suggested framework, where X_{io} represents the initial vector for client sites using the M-ABC algorithm; X_{ri} is the randomly chosen local solution; the candidate solutions of employed, onlooker, and scout bees are represented as C_{en}, C_{on}, and C_{sn}, respectively; Fit[n] is the fitness function; B is the size of the local batch at HSP clients; ω_o is the initial model disseminated by the HSP global orchestrator; similarity function is represented as $c\ (w_{jl}, \theta_i)$; and the decision function based on RB-SVM for the heart disease dataset d is represented by $D_F\ (d)$.

Table 1. Brief description of the symbols utilized in our proposed framework.

Symbol	Brief Description
X_{io}	Initial vector for MABC at client sites
X_{ri}	Local solution chosen randomly
C_{en}	Employed bee's candidate solution
C_{on}	Candidate solution from onlooker bee
C_{sn}	Candidate solution obtained by scout bee
Fit[n]	Fitness function
N	Number of HSP clients
B	Local minibatch at every HSP client
ω_o	Initial model by HSP global orchestrator
$c\ (w_{jl}, \theta_i)$	Similarity function
θ_i	Gaussian mean
w_{jl}	Weight of lth neuron on dataset j in MABC
E	Local epochs
η	Learning rate
ω_N	Model of Nth HSP client
d	Input dataset to RB-SVM
K_{F}	Kernel function
$D_F\ (d)$	Decision function on dataset d in RB-SVM
m_r	Margin function
RB_{F}	Radial basis function

In the FL-based HSP environment, we have a small number of HSP client devices and the data for independent model training on each device are not sufficient; an FedMA approach is more suitable. Our idea is to train multiple models independently on every HSP client and then average their predictions to produce a final prediction. This can lead to a better performance by reducing overfitting and increasing the diversity of the models. We propose an FL-based framework for privacy-aware prediction of the heart disease. Our proposed framework is constituted for the HSP client- and server-end. For the HSP client, the hybrid model of the RB-SVM with M-ABC for optimal classification and feature selection is proposed. For the HSP server, the FedMA is proposed to overcome the issues of HSP's central model updation and communication efficiency.

4.1. M-ABC-Based Feature Selection

At the HSP client-end, the M-ABC optimizer is used to choose features in the best way possible. The implementation of the M-ABC is completed in the four phases and each phase is described as follows:

4.1.1. Phase-I

Every client site of HSP is initialized by X_{io} vector. This initialization is achieved by using Equation (1):

$$X_{io} = l_i + rand\ (0,1) * (u_i - l_i) \quad (1)$$

where the u_i and l_i represent the upper and lower bounds of the parameters, respectively.

4.1.2. Phase-II (Searching of Candidate Solution by Employed Bee)

Using Equation (2), the bee searches the local HSP clients for new candidate solutions during this phase. The random integer generated by the function τ_{ni} falls between [−1 and 1], and the local random solution is represented by X_{ri}.

$$C_{en}[i] = X_{io} + \tau_{ni} * (X_{io} - X_{ri}) \quad (2)$$

The fitness function Fit[n] using Equation (3) determines the fitness of a new candidate solution, and if the fitness value is high, the solution is memorized.

$$\mathrm{Fit}[n] = \begin{cases} \frac{1}{1+F_{obj}}, \text{ if } F_{obj} \geq 0, \\ 1 + \mathrm{abs}(F_{obj}), \text{ if } F_{obj} < 0. \end{cases} \quad (3)$$

4.1.3. Phase-III (Onlooker Bee's Candidate Solution)

Employed bees present their potential solution to the onlooker bee, who then makes a probabilistic decision C_{on} using Equation (4).

$$C_{on}[i] = \frac{F_n[i]\ (X_n)}{\sum_{i=1}^{m} (F_m)(X_n)} \quad (4)$$

4.1.4. Phase-IV (Scout Bee with Firefly)

By picking a C_{sn} solution using the firefly process as shown in the equation below, the scout bee ensures that the new solution is evaluated in Equation (5). An employed bee becomes a scout bee if it fails to improve its solution within a defined time range.

$$C_{sn}[i] = C_{sn}[i] + e^{-r_i{}^2}(C_{sn0}[i] - C_{sn}[i]) + (rand\ (0,1) - 0.5) \quad (5)$$

4.2. Classification Based on SVM

For a non-linear classification problem with a multidimensional set of features, the decision function using Equation (6) in terms of kernel function K_F (d, d_j) for the input dataset d, m_r as the margin, and weight represented as βj, can be written as:

$$D_F\ (d) = \sum_{d=1}^{n} \beta j.\ K_F\ (d,\ d_j) + m_r \quad (6)$$

To solve the heart disease (non-linear discrete) classification problem with a feature set of a high-dimension, kernel function is modified to be the radial basis function (RBF) RB_F $(d, d_j) = e^{(-\gamma \mid [d - dj] \mid \hat{} 2)}$. Therefore, the decision function defined in the above Equation (6) is modified and is computed using Equation (7). In our proposed framework, the RB-SVM classifier is implemented at the HSP client side.

$$D_F\ (d) = \sum_{d=1}^{n} \beta j.\ RB_F\ (\mathrm{d},\ \mathrm{dj}) + m_r \quad (7)$$

4.3. Discussion on Proposed Framework

In this section, we present a full review of our proposed framework comprised of the hybrid MABC-SVM with FedMA-based technique for the effective prediction of CVD. Our suggested system model is depicted in Figures 1 and 2. Our suggested system consists of heart disease data-gathering equipment housed within a healthcare facility. Initially, the HSP global model orchestrator distributes a global model to HSP clients. The HSP clients, upon reception of this model, perform classification and optimal feature selection using our proposed MABC-SVM technique, and then perform the local training. The HSP client nodes send their updated local model towards the HSP central orchestrator. A new global model is generated using the FedMA after receiving repeated updates for local models, and it is distributed among the HSP clients. According to our proposed approach, the privacy issues are addressed as all of the CVD data never left the HSP client node, prediction accuracy is raised, and classification mistakes are decreased. Algorithm 1 below shows how our proposed framework functions.

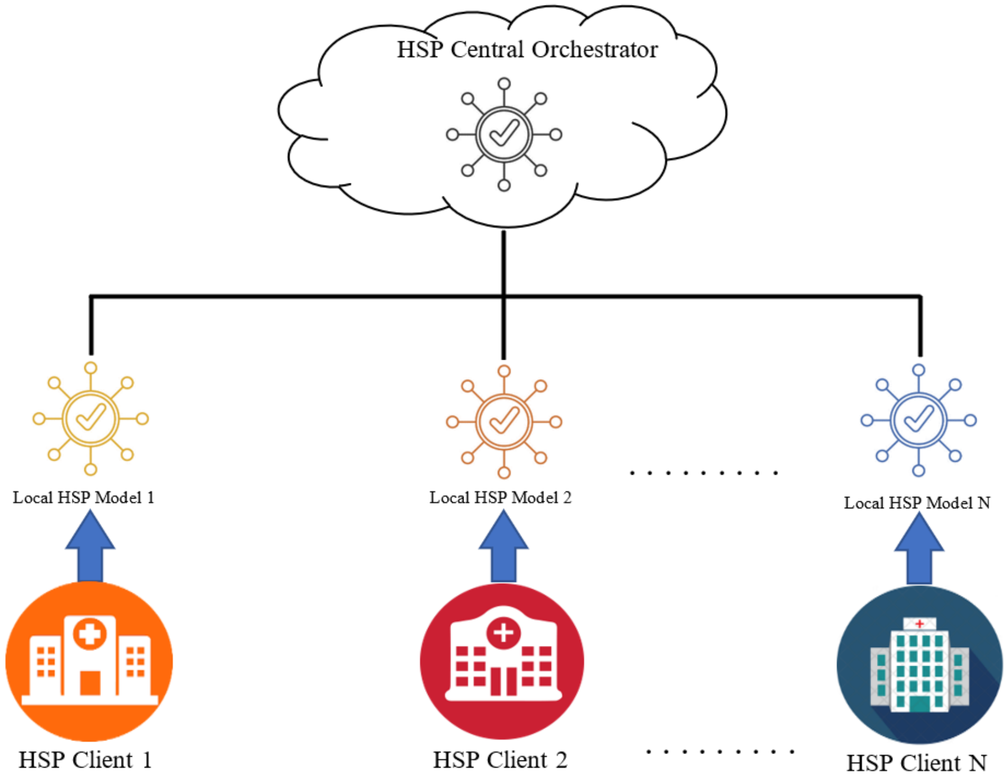

Figure 1. Overview of proposed hybrid federated learning framework.

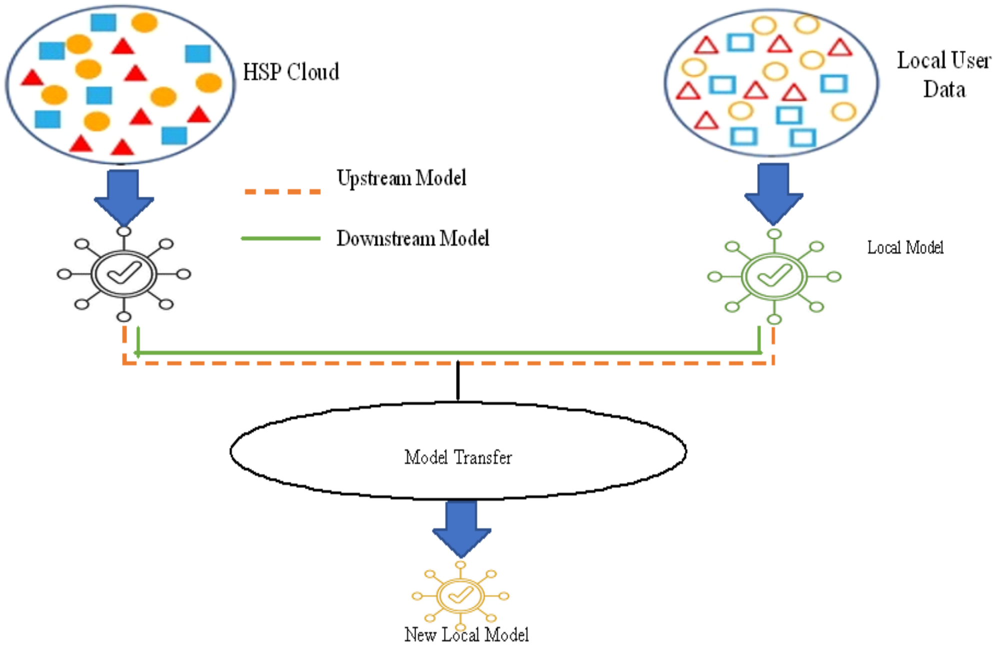

Figure 2. Illustration of model computation at HSP global model orchestrator and user ends.

The proposed FedMA-based framework obtains the maximum aposteriori estimate (MAE) using the Bernoulli process described below:

$$\min_{\{\pi^j_{li}\}} \sum_{i=1}^{L} \sum_{j,l} \min_{\theta_i} \pi^j_{li}.\mathrm{c}(w_{jl})\mathrm{s.t.} \sum_i \pi^j_{li} = 1 \forall j, l;\ \sum_l \pi^j_{li} = 1 \forall\ i,\ j. \tag{8}$$

where the w_{jl} is lth neuron of the dataset j and an appropriate function of similarity is *c(..)*. The posterior probability of $c\ (w_{jl}, \theta_i)$ is computed on the jth client neuron l and mean Gaussian θ_i. The total neurons in the federated model can gradually increase in accordance to the sizes of the HSP client models because our suggested inference approach is not reliant on parameters. Our proposed framework is designed for the HSP clients and cloud sites, and it is executed in the following stages:

1. Stage-I (initial): An initial global model $\omega_{\,o}$ is disseminated to every HSP client user HC_N. After obtaining this initial model, the HSP client is initiated for initial feature selection using X_{i0}.
2. Stage-II (HSP clients): The client nodes will perform feature selection and classification of each fragmented local data of size β using a hybrid MABC with the RB-SVM technique. The updated weights of the local solution are returned to the HSP global orchestrator from every HSP client.
3. Stage-III (HSP global orchestrator): Upon reception of the weights from every HSP client, it performs the matched averaging and obtains an updated weight ω_N for the current round of communication.
4. Stage-IV (finalization at HSP global orchestrator): The updated weights ω_N are computed until there is no evolution in the HSP client models.

Algorithm 1: Proposed hybrid FL-based framework for heart disease prediction

Input: CVD Data from HSP clients $\{HC_1, HC_2, ---, HC_N\}$
Output: Privacy aware model for heart disease at HSP client user ω_N
// Computation at the HSP global orchestrator:

1: Initialize with global model ω_o
2: **for** each round $i = 1, 2, \ldots$ **do**

i) $m \leftarrow max\ (N, 1)$
ii) $S[t] \leftarrow$ (m is selected randomly for HSP clients)

3: **do in parallel for** each client $N \in S[t]$

(i) Compute inference method using Equation (8) with $(\{N, C_n, \omega_m\})$

(ii) $\omega_N \leftarrow \frac{1}{N} \sum_{k=1}^{N} \omega_m^N \prod_m^N$

(iii) $\omega_{m+1} \leftarrow \prod_m^N \omega_N$ //next weights permutation

4: Disseminate ω_N among the HSP clients
5: Repeat until no evolution found in client models

// Computation at the HSP Client End (N, ω)**:**

1: **foreach** client in N
  (i) $\beta \leftarrow$ (fragment the local data to β size each into P_N groups)
  (ii) Calculate C_n using MABC with Equations (2), (4) and (5)
  (iii) Perform the decision classification using RB-SVM classifier with Equation (7)

2: **for** every local $i = 1 \ldots E$ epochs **do**

(i) **for** $b \in \beta$ groups **do**
  (a) Perform gradient descent using $(\omega; b)$

3: send back ω to the HSP global orchestrator

5. Experimental Evaluation and Validation

5.1. Simulation Setup

We use the Python environment (*PyTorch*) on a system with an Intel ® Core TM i7 @ 4 GHz and 16 GB RAM to run a simulation with 5000 communication rounds to evaluate the performance of the proposed framework. For standard FL algorithms at the client end, we implement the SVM classifier. Hence, the standard FL algorithms are versions of the SVM such as FedAvg-SVM, and FedMA-SVM, and we also develop an upgraded version of the vanilla FedMA compromising of a genetic algorithm (GA) and SVM as FedMA with GA-SVM. The effectiveness of our framework for heart disease is assessed and compared with state-of-the-art FedAvg, FedMA, and FedMA with GA-SVM approaches in terms of prediction accuracy, time to attain the accuracy, communication efficiency, and influence of the local epoch on accuracy. We consider the number of HSP client nodes to be five and one HSP server node. However, this proposed framework can be scaled-up for the HSP client nodes.

5.2. Dataset Description

Utilizing the combined dataset of five heart illness datasets, this dataset combines over eleven common features from the datasets of Cleveland, Stalog, Hungary, Long Beach, and Switzerland [43]. This dataset is used for the prediction of CVD, and it consists of various parameters for CVD. The dataset has records of CVD patients recorded using eleven heart disease features. The eleven CVD features of this dataset include resting blood pressure, cholesterol serum, chest pain, max heart rate, depression level, resting electrocardiogram, angina-induced by exercise, fasting blood sugar, ST slope, age, and sex. We train and evaluate our suggested framework on this combined dataset (this dataset is available at https://www.kaggle.com/fedesoriano/heart-failure-prediction, accessed on 8 December

2022). There are 918 entries in the combined dataset collection, along with 76 characteristics in each dataset. Table 2 provides an illustration of the dataset's complete description. The many risks for developing heart disease are included in this table along with their descriptions and encoded values. Our suggested approach uses the encoded values as its input. For the experimentation of this dataset using our proposed framework, the control group refers to the group of patients who do not have heart disease (as determined by the target column of Table 2). The patients with heart failure are considered the experimental group. The target column in the dataset is used to distinguish between the two groups, with 0 indicating no heart failure and 1 indicating heart failure.

Table 2. Thorough description of the combined dataset.

S#	Feature	Explanation	Unit	Coded Values
1	Resting blood pressure (Rt_Bp)	In mmHg	Integer	Low Level = Below 120 = −1 Normal Level = 120–139 = 0 High Level = Above 139 = 1
2	Cholesterol serum (Cl_S)	In mg/dL	Integer	<200 mg/dL = Low = −1 200–239 mg/dL = Normal = 0 >240 mg/dL = High = 1
3	Chest pain (C_P)	Type of chest pain	String	Angina Typical (AT) = 2 Asymptomatic (AS) = 1 Angina Atypical (ATA) = 0 Non-Angina (NA) = −1
4	Max heart rate (MHR)	Maximum achieved heart rate in bpm	Integer	<69 bpm = Low = −1 70–90 bpm = Normal = 0 >91 bpm = High = 1
5	Depression level (Dp_L)	Old peak in ST (numeric value measured for depression level)	Float	<0.5 mm = Normal = 0 >0.5 mm = High = 1
6	Resting electro-cardiogram (Rt_ECG)	Normal, ST T, or LVH	String	LVH = 2 ST T = 1 Normal = 0
7	Angina induced by exercise (AI_bE)	Yes or No	String	Yes = 1 No = 0
8	Fasting blood sugar (F_BS)	>120 mg/dL	Integer	True = 1 False = 0
9	ST slope (ST_S)	Peak exercise slope	String	Up = 2 Flat = 1 Down = 0
10	Age (A)	Age in years	Integer	>77 = 2, 64–77 = 1, 47–63 = 0, 35–46 = −1, <35 = −2
11	Sex (S)	Female and Male	String	Female = 0, Male = 1
12	Target (heart disease)	Yes or No	Integer	Yes = 1, No = 0

5.3. Results and Discussion

In the FL-based HSP environment, the communication rounds refer to the number of times the HSP client model parameters are exchanged during the training process. If the large number of rounds are consumed by an FL model, then it will also increase the communication overhead and computational cost. Therefore, the number of communication rounds is a key aspect in the FL system's overall performance and efficiency. The impact of communication rounds on the algorithm's accuracy in making predictions on the combined dataset is seen in Figure 3. Our proposed framework reaches 93.8% accuracy within the 4500 rounds of communication, which is better than the existing FedAvg-SVM, FedMA-SVM, and FedMA with GA-SVM algorithms. Since the proposed MABC-RB-SVM method employs the hybrid of the MABC optimizer and RB-SVM classifier for optimal feature selection and classification at HSP clients, and for the HSP global orchestrator, we deploy the FedMA which permits our overall model to accomplish better accuracy in a lesser number of communication rounds than the existing FL algorithms. In the FedAvg-SVM and FedMA-SVM, the HSP overall model performs the simple averaging and matched averaging, respectively, on the simple SVM kernel at their client model algorithm which results in consuming higher communication rounds. The learning rate is increased in the GA-SVM with FedMA, but the convergence is consumed more by the classification and feature selection. Consequently, the proposed hybrid framework accomplishes better accuracy and reduces the amount of communication rounds used.

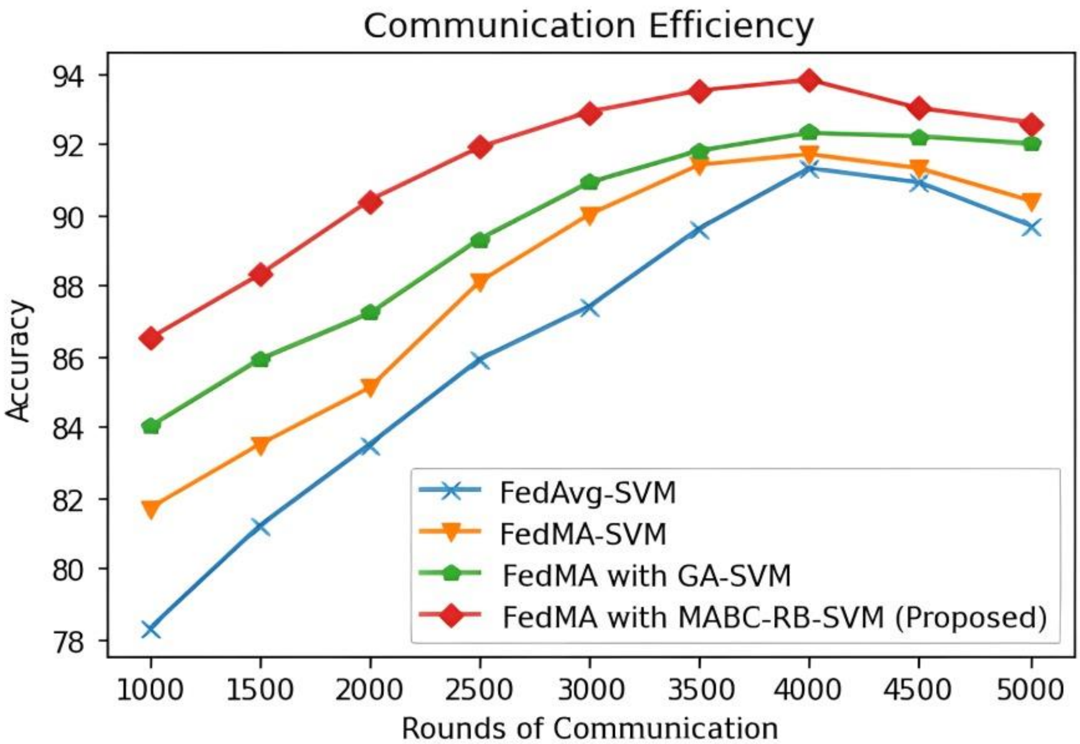

Figure 3. Comparison of convergence rate with prediction accuracy.

We evaluate and vary the local epochs E from 10 to 160, to examine the impact of local epochs on the prediction accuracy of the proposed hybrid technique and existing FL algorithms. The accuracy test on each E of the proposed hybrid framework, FedAvg-SVM, FedMA-SVM, and hybrid of GA-SVM with FedMA, is reviewed and compared. Figure 4 shows the outcome of this test. The findings show that the suggested framework can train for a longer period and supports a higher rate of convergence because it produces a better HSP global model on the local model with a higher model quality. This is because our suggested technique makes use of the RB-SVM and MABC methods at the client side of HSP. The accuracy of traditional FL algorithms, such as FedAvg and FedMA, tends to decrease with time due to the lack of an optimizer. However, in the case of the GA-SVM with FedMA, the accuracy does not decrease much after 100 local epochs, which is attributable to the GA algorithm. This result demonstrates that if user sites implement our recommended structure, they are free to train the local users of their model indefinitely.

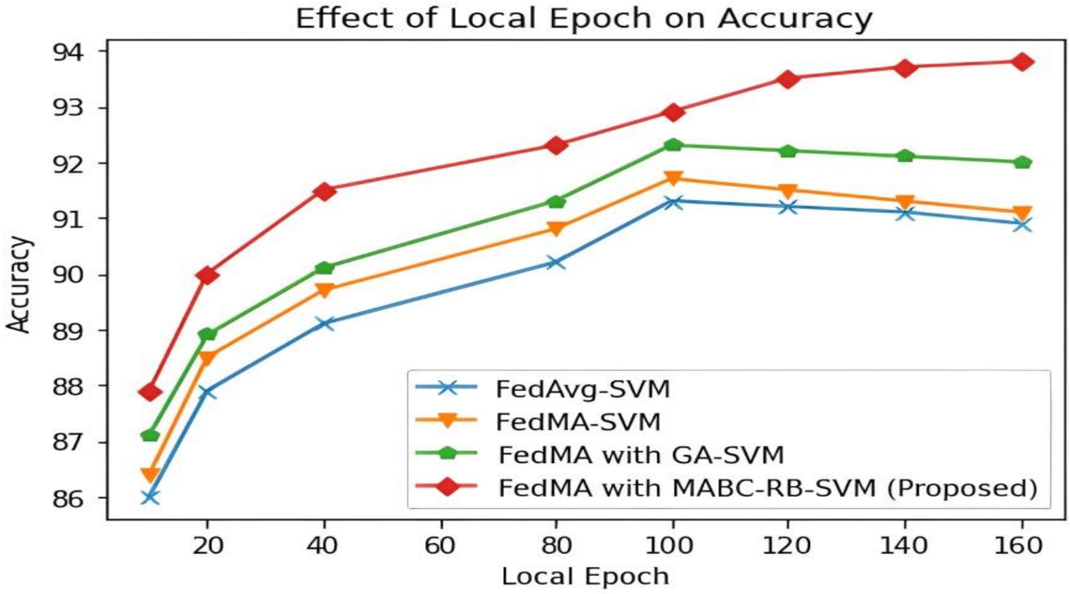

Figure 4. Comparison of the effect of local epoch on the prediction accuracy.

For the influence of prediction accuracy on the utilized communication volume, we examine and compare the performances of the FedAvg-SVM, Fed-MA-SVM, GA-SVM with FedMA, and our suggested approach. We vary the volume of communication (in Gigabytes) for this assessment as {0.6, 1.2, 1.8, 2.4, 3.0, 3.6, 4.2, 4.8, 5.4, 6.0} and record the prediction accuracy of each approach, as shown in Figure 5. The recorded results show that our hybrid approach outperforms traditional FL techniques and FedMA with the GA-SVM in terms of accuracy at both low and large communication volumes. Furthermore, Figure 6 depicts a comparison of the extent of communication necessary to achieve the various target prediction accuracy of the algorithms (70%, 75%, 80%, 85%, and 90%). The GA-SVM outperforms our suggested approach for a lower target accuracy of 70% and 75%. However, when compared to existing FL algorithms, our proposed approach consumes 15–25% less communication size (in GB) for improved target accuracy.

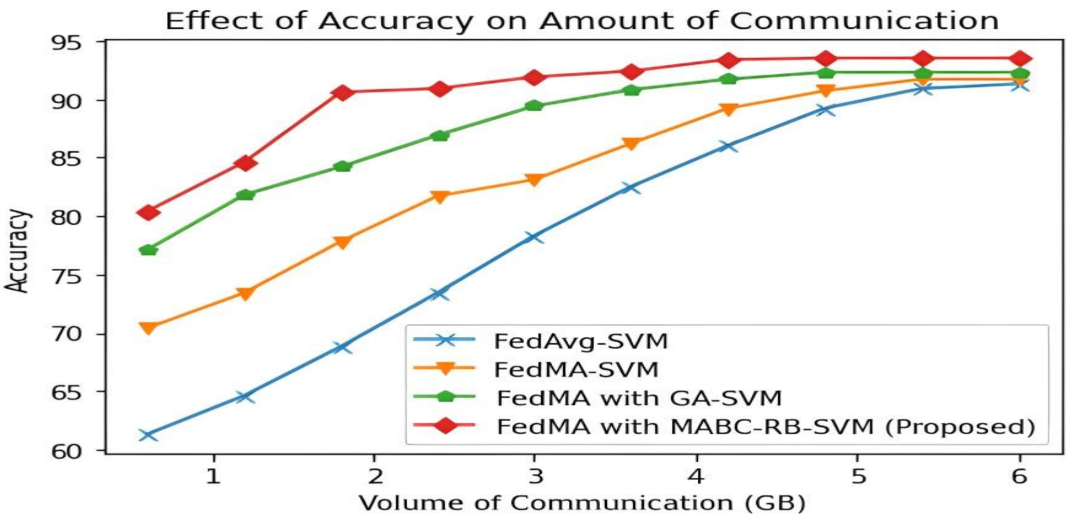

Figure 5. Analysis of prediction accuracy on the volume of communication.

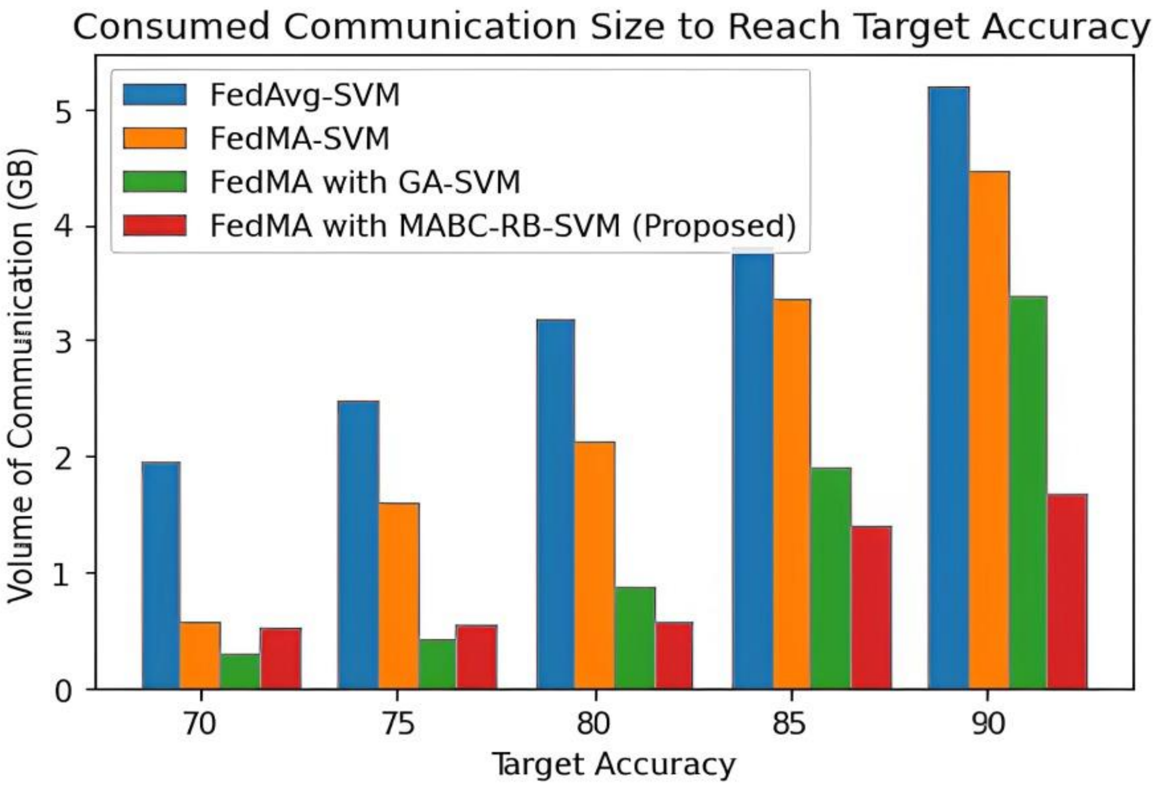

Figure 6. Comparison of consumed communication volume to reach target prediction accuracies.

Performance metrics such as accuracy, precision, classification error, f-measure, specificity, sensitivity, and the number of rounds required to achieve the highest accuracy are assessed for the performance efficiency comparison of the proposed framework with the FedAvg-SVM, FedMA-SVM, and GA-SVM with FedMA. Accuracy in machine learning refers to the proportion of all available examples that yield the right predictions. The fraction of accurate positive instance predictions is what is referred to as precision. Classification errors are defined as the inaccuracies or proportions of mistakes that are readily available in the instance. Three performance indicators are used to identify important heart disease symptoms. This makes it easier to comprehend how different groups behave and enables better feature selection. The results of these parameters are displayed in Tables 3 and 4. The created GA-SVM with FedMA for the heart disease dataset and baseline FL approaches are compared to our proposed framework, which achieves greater target accuracy in less cycles. The number of rounds in our suggested technique is reduced by 37% when compared to existing methods, as shown in Table 3, since our proposed model's learning rate grows quickly after each round, leading to fewer rounds. Table 3 demonstrates that the proposed framework performs better on the heart disease dataset than FL state-of-the-art methods in terms of prediction accuracy (93.8%), precision (94.2%), sensitivity (96.6), and specificity (81.8), because of the proposed model's improved learning rate that increases with each communication round with a smaller minibatch size. Our proposed framework is therefore more equipped to provide increased heart disease prediction accuracy while maintaining privacy when compared to existing baseline FL techniques. In addition, the MABC optimization technique for feature selection and the RB-SVM classification in our proposed framework result in decreased classification errors, resulting in a classification error of 11.9 for our recommended method.

Table 3. Consumed algorithm time for the highest accuracy of the model.

Techniques	Max. Accuracy Achieved	# Of Rounds to Reach 91%	Difference
FedAvg-SVM	91.3	3810	–
FedMA-SVM	91.7	3425	10.1%
FedMA with GA-SVM	92.3	3046	20.1%
FedMA with MABC-RB-SVM (Proposed)	**93.8**	**2408**	**37.8%**

Table 4. Comparison of performance on features of the dataset.

Techniques	Accuracy	F-Measure	Precision	Classification Error	Sensitivity	Specificity
FedAvg-SVM	91.3	87.3	92.3	20.4	85.3	59.5
FedMA-SVM	91.7	88.4	90.1	18.6	89.5	72.5
FedMA with GA-SVM	92.3	89.6	93.7	13.3	91.9	78.8
FedMA with MABC-RB-SVM (Proposed)	**93.8**	**90.1**	**94.2**	**11.9**	**96.6**	**81.8**

6. Conclusions and Future Work

For the objective of early illness detection and treatment, a health service provider (HSP) system to collect patient data in real time has been created. Intelligent healthcare systems can move quickly to save many lives, especially when a patient is in a remote place without access to medical treatment. It is challenging to predict survival in patients with cardiac disease. Due to privacy and security concerns, it is hard to exchange user data when it comes to healthcare systems. In this paper, we proposed a hybrid federated learning framework for improved heart disease prediction and to address privacy issues in the healthcare system. In order to enhance heart disease prediction, the proposed framework combines MABC with RB-SVM feature optimization and classification techniques at the HSP's client node, while FedMA is used at the HSP global orchestrator to solve communication efficacy and privacy problems in the healthcare system. The main goal of this research is to shorten training time and improve communication efficiency while improving the prediction accuracy of heart disease. We evaluated and compared the performance in terms of several model prediction-based metrics and communication efficiency with the baseline FedMA, FedMA, and a developed upgraded version of FedMA using a GA-SVM optimizer and classifier algorithms in order to ensure the accuracy and validity of our proposed framework. Performance metrics including prediction accuracy, classification error, sensitivity, precision, and communication efficiency all showed a considerable improvement under the suggested paradigm. Our findings indicated that the suggested strategies produce outcomes with 1.5% greater accuracy, 1.6% lower classification error, 4.7% higher sensitivity, and 17.7% fewer rounds needed to reach the greatest degree of accuracy. In the future, we will focus on the rehabilitation and treatment of several additional serious illnesses including Parkinson's, diabetes, liver cancer, skin cancer, and breast cancer.

Author Contributions: Conceptualization, M.M.Y. and M.N.; methodology, M.M.Y., M.A.K. and A.A.-R.; software, M.M.Y. and S.Q.; validation, M.M.Y., M.N., A.A.-R. and M.A.K.; formal analysis, M.M.Y. and M.N.; investigation, M.M.Y. and S.Q.; resources, M.M.Y. and M.A.K.; data curation, M.M.Y. and M.N.; writing—original draft preparation, M.M.Y., M.N. and M.A.K.; writing—review and editing, M.M.Y., M.N., S.Q. and M.A.K.; visualization, M.M.Y. and M.A.K.; supervision, M.N.; project administration, M.N., S.Q., A.A.-R. and M.A.K.; funding acquisition, A.A.-R. All authors have read and agreed to the published version of the manuscript.

Funding: Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R235), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: We ran simulations to see how well the proposed approach performed. Any questions concerning the study in this publication are welcome and can be directed to the lead author (Muhammad Mateen Yaqoob) upon request.

Acknowledgments: The authors sincerely appreciate the support from Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R235), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Turjman, F.A.; Nawaz, M.H.; Uluser, U.D. Intelligence in the Internet of Medical Things era: A systematic review of current and future trends. *Comput. Commun.* **2020**, *150*, 644–660. [CrossRef]
2. Dash, S.; Shakyawar, S.K.; Sharma, M.; Kaushik, S. Big data in healthcare: Management, analysis and future prospects. *J. Big Data* **2019**, *6*, 54. [CrossRef]
3. Watkins, D.A.; Beaton, A.Z.; Carapetis, J.R.; Karthikeyan, G.; Mayosi, B.M.; Wyber, R.; Yacoub, M.H.; Zühlke, L.J. Rheumatic heart disease worldwide: JACC scientific expert panel. *J. Am. Coll. Cardiol.* **2018**, *72*, 1397–1416. [CrossRef] [PubMed]
4. Mohan, S.; Thirumalai, C.; Srivastava, G. Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques. *IEEE Access* **2019**, *7*, 81542–81554. [CrossRef]
5. Li, J.P.; Haq, A.U.; Din, S.U.; Khan, J.; Khan, A.; Saboor, A. Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare. *IEEE Access* **2020**, *8*, 107562–107582. [CrossRef]
6. Voigt, P.; dem Bussche, A.V. Scope of application of the GDPR. In *The EU General Data Protection Regulation*; Springer: Cham, Switzerland, 2017; pp. 9–30.
7. Wagner, J. China's Cybersecurity Law: What You Need to Know. The Diplomat. 2017. Available online: https://thediplomat.com/2017/06/chinas-cybersecurity-law-what-you-need-to-know/ (accessed on 10 October 2022).
8. de la Torre, L. A Guide to the California Consumer Privacy Act of 2018. 2018. Available online: http://dx.doi.org/10.2139/ssrn.3275571 (accessed on 10 October 2022).
9. McMahan, B.; Ramage, D. Federated Learning: Collaborative Machine Learning without Centralized Training Data. Google AI Blog 2017. Available online: https://ai.googleblog.com/2017/04/federated-learning-collaborative.html (accessed on 11 October 2022).
10. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Areas, B.A.Y. Communication-Efficient Learning of Deep Networks from Decentralized Data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL, USA, 20–22 April 2017; pp. 1273–1282.
11. Yaqoob, M.M.; Nazir, M.; Yousafzai, A.; Khan, M.A.; Shaikh, A.A.; Algarni, A.D.; Elmannai, H. Modified Artificial Bee Colony Based Feature Optimized Federated Learning for Heart Disease Diagnosis in Healthcare. *Appl. Sci.* **2022**, *12*, 12080. [CrossRef]
12. Xu, X.; Liu, W.; Zhang, Y.; Zhang, X.; Dou, W.; Qi, L.; Bhuiyan, M.Z.A. Psdf: Privacy-aware iov service deployment with federated learning in cloud-edge computing. *ACM Trans. Intell. Syst. Technol. (TIST)* **2022**, *13*, 70. [CrossRef]
13. Zhang, X.; Hu, M.; Xia, J.; Wei, T.; Chen, M.; Hu, S. Efficient Federated Learning for Cloud-Based AIoT Applications. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2021**, *40*, 2211–2223. [CrossRef]
14. Yang, J.; Zheng, J.; Zhang, Z.; Chen, Q.; Wong, D.S.; Li, Y. Security of federated learning for cloud-edge intelligence collaborative computing. *Int. J. Intell. Syst.* **2022**, *37*, 9290–9308. [CrossRef]
15. Elaziz, M.A.; Xiong, S.; Jayasena, K.P.N.; Li, L. Task scheduling in cloud computing based on hybrid moth search algorithm and differential evolution. *Knowl.-Based Syst.* **2019**, *169*, 39–52. [CrossRef]
16. Yang, Y.; Chen, H.; Heidari, A.A.; Gandomi, A.H. Hunger games search: Visions, conception, implementation, deep analysis, perspectives, and towards performance shifts. *Expert Syst. Appl.* **2021**, *177*, 114864. [CrossRef]
17. Panniem, A.; Puphasuk, P. A Modified Artificial Bee Colony Algorithm with Firefly Algorithm Strategy for Continuous Optimization Problems. *J. Appl. Math.* **2018**, *2018*, 1237823. [CrossRef]
18. Chen, H.L.; Yang, B.; Liu, J.; Liu, D.-Y. A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. *Expert Syst. Appl.* **2011**, *38*, 9014–9022. [CrossRef]
19. Yadav, D.P.; Saini, P.; Mittal, P. Feature Optimization Based Heart Disease Prediction using Machine Learning. In Proceedings of the 2021 5th IEEE International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 22–23 October 2021; pp. 1–5.
20. Abdulrahman, S.; Tout, H.; Ould-Slimane, H.; Mourad, A.; Talhi, C.; Guizani, M. A Survey on Federated Learning: The Journey From Centralized to Distributed On-Site Learning and Beyond. *IEEE Internet Things J.* **2021**, *8*, 5476–5497. [CrossRef]

21. Brisimi, T.S.; Chen, R.; Mela, T.; Olshevsky, A.; Paschalidis, I.C.; Shi, W. Federated learning of predictive models from federated electronic health records. *Int. J. Med. Inform.* **2018**, *112*, 59–67. [CrossRef]
22. Nilsson, A.; Smith, S.; Ulm, G.; Gustavsson, E.; Jirstrand, M. A performance evaluation of federated learning algorithms. In Proceedings of the Second Workshop on Distributed Infrastructures for Deep Learning (DIDL), Rennes France, 10–11 December 2018; ACM: New York, NY, USA, 2018; pp. 1–8.
23. Wang, H.; Yurochkin, M.; Sun, Y.; Papailiopoulos, D.; Khazaeni, Y. Federated learning with matched averaging. *arXiv* **2020**, arXiv:2002.06440.
24. Aledhari, M.; Razzak, R.; Parizi, R.; Saeed, F. Federated Learning: A Survey on Enabling Technologies, Protocols, and Applications. *IEEE Access* **2020**, *8*, 140699–140725. [CrossRef]
25. Ma, Z.; Mengying, Z.; Cai, X.; Jia, Z. Fast-convergent federated learning with class-weighted aggregation. *J. Syst. Archit.* **2021**, *117*, 102125. [CrossRef]
26. Salam, M.A.; Taha, S.; Ramadan, M. COVID-19 detection using federated machine learning. *PLoS ONE* **2021**, *16*, e0252573.
27. Cheng, W.; Ou, W.; Yin, X.; Yan, W.; Liu, D.; Liu, C. A Privacy-Protection Model for Patients. *Secur. Commun. Netw.* **2020**, *2020*, 6647562. [CrossRef]
28. Fang, L.; Liu, X.; Su, X.; Ye, J.; Dobson, S.; Hui, P.; Tarkoma, S. Bayesian Inference Federated Learning for Heart Rate Prediction. In Proceedings of the International Conference on Wireless Mobile Communication and Healthcare, Virtual Event, 19 November 2020; Springer: Cham, Switzerland, 2020; pp. 116–130.
29. Babar, M.; Khan, M.; Din, A.; Ali, F.; Habib, U.; Kwak, K.S. Intelligent Computation Offloading for IoT Applications in Scalable Edge Computing Using Artificial Bee Colony Optimization. *Complexity* **2021**, *2021*, 5563531. [CrossRef]
30. Nguyen, H.T.; Sehwag, V.; Hosseinalipour, S.; Brinton, C.; Chiang, M.; Poor, H.V. Fast-Convergent Federated Learning. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 201–218. [CrossRef]
31. Manimurugan, S.; Almutairi, S.; Aborokbah, M.; Narmatha, C.; Ganesan, S.; Chilamkurti, N.; Alzaheb, R.A.; Almoamari, H. Two-Stage Classification Model for the Prediction of Heart Disease Using IoMT and Artificial Intelligence. *Sensors* **2022**, *22*, 476. [CrossRef] [PubMed]
32. Yuan, X.; Chen, J.; Zhang, K.; Wu, Y.; Yang, T. A Stable AI-Based Binary and Multiple Class Heart Disease Prediction Model for IoMT. *IEEE Trans. Ind. Inform.* **2022**, *18*, 2032–2040. [CrossRef]
33. Khan, M.A.; Algarni, F. A Healthcare Monitoring System for the Diagnosis of Heart Disease in the IoMT Cloud Environment Using MSSO-ANFIS. *IEEE Access* **2020**, *8*, 122259–122269. [CrossRef]
34. Chhabra, A.; Singh, G.; Kahlon, K.S. Multi-criteria HPC task scheduling on IaaS cloud infrastructures using meta-heuristics. *Clust. Comput.* **2021**, *24*, 885–918. [CrossRef]
35. Li, C.; Hu, X.; Zhang, L. The IoT-based heart disease monitoring system for pervasive healthcare service. *Procedia Comput. Sci.* **2017**, *112*, 2328–2334. [CrossRef]
36. Khan, M.A. An IoT Framework for Heart Disease Prediction Based on MDCNN Classifier. *IEEE Access* **2020**, *8*, 34717–34727. [CrossRef]
37. Sarmah, S.S. An Efficient IoT-Based Patient Monitoring and Heart Disease Prediction System Using Deep Learning Modified Neural Network. *IEEE Access* **2020**, *8*, 135784–135797. [CrossRef]
38. Makhadmeh, Z.A.; Tolba, A. Utilizing IoT wearable medical device for heart disease prediction using higher order Boltzmann model: A classification approach. *Measurement* **2019**, *147*, 106815. [CrossRef]
39. Ganesan, M.; Sivakumar, N. IoT based heart disease prediction and diagnosis model for healthcare using machine learning models. In Proceedings of the 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 29–30 March 2019; pp. 1–5.
40. Albahri, A.S.; Zaidan, A.A.; Albahri, O.S.; Zaidan, B.B.; Alamoodi, A.H.; Shareef, A.H.; Alwan, J.K.; Hamid, R.A.; Aljbory, M.T.; Jasim, A.N.; et al. Development of IoT-based mhealth framework for various cases of heart disease patients. *Health Technol.* **2021**, *11*, 1013–1033. [CrossRef]
41. Gupta, A.; Yadav, S.; Shahid, S.; Venkanna, U. HeartCare: IoT Based Heart Disease Prediction System. In Proceedings of the 2019 International Conference on Information Technology (ICIT), Bhubaneswar, India, 19–21 December 2019; pp. 88–93.
42. Jabeen, F.; Maqsood, M.; Ghanzafar, M.A.; Adil, F.; Khan, S.; Khan, M.F.; Mehmood, I. An IoT based efficient hybrid recommender system for cardiovascular disease. *Peer-to-Peer Netw. Appl.* **2019**, *12*, 1263–1276. [CrossRef]
43. Fedesoriano. Heart Failure Prediction Dataset. Available online: https://www.kaggle.com/fedesoriano/heart-failure-prediction (accessed on 28 November 2022).