[1]Sulakshana Malwade

[2]Asha R. Sanap

[3]Yogesh Sudam Gite

[4]Pranjal Pandit

[5]Yogesh J. Gaikwad

[6]Dr. Amruta Mhatre

Predicting Heart Diseases in IoT-Based Electronic Health Records: A Federated Learning Approach

Abstract: - Predicting heart diseases is important for finding them early and treating them effectively. We present a shared learning method for predicting heart diseases using IoT-based electronic health records (EHRs) in this work. Federated learning lets many autonomous IoT devices work together to train a model, while protecting the safety and security of the data. Proposed method uses the fact that IoT devices are spread out to train a global model for predicting heart disease without putting private EHR data in one place. With the EHR data, each IoT device learns a model locally and only sends model changes to a central computer. The computer takes all of these changes and improves the world model. This model is then sent back to the IoT devices to be improved even more. This looping process makes sure that the world model keeps getting better while keeping data private. The proposed method tested by using a real-world collection of EHRs from IoT devices in trials. We looked at how well our shared learning method worked compared to more standard centralized learning methods. Our results show that the pooled learning method makes predictions that are as good as or better than the other methods while protecting data privacy. It also looked at how different IoT device properties, like the amount of data they send and receive and their processing power, affect the shared learning process. It is discovered that devices with more processing power and more data add more to the improvement of the global model. This shows how important it is to choose the right devices in shared learning systems. The paper study shows that pooled learning can be used to predict heart diseases in IoT-based EHRs and that it works well. Our method uses the ability of IoT devices to work together to make accurate predictions while protecting data privacy. This makes it suitable for use in real-life healthcare situations.

Keywords: Heart diseases, Federated learning, IoT-based electronic health records, Data privacy, Healthcare applications

I. INTRODUCTION

Heart diseases are the top cause of death in the world. In 2019, they will be responsible for 17.9 million deaths, or 32% of all deaths in the world. Early detection and treatment are very important for handling heart illnesses well and lowering death rates. The Internet of Things (IoT) has changed healthcare in recent years by letting doctors keep an eye on patients' health all the time through smart tech and monitors. Electronic health records (EHRs) and other types of data that these devices create can be used to make models that can predict heart illnesses [1]. Adding Internet of Things (IoT) devices to healthcare systems could make care better by giving doctors more real-time

[1] Assistant Professor, Department of Computer Engineering, Dr. Vishwanath Karad World Peace University, Pune, India

sulakshana.malwade@mitwpu.edu.in

2Assistant Professor, Program - ECE, Department of Polytechnic, Dr. Vishwanath Karad World Peace University, Pune, India

asha.sanap@mitwpu.edu.in

3Assistant Professor, G.H. Raisoni College of Engg and Management, Pune, India

yogeshgite3917@gmail.com

4Assistant Professor, Department of Computer Science and Engineering -Artificial Intelligence, Vishwakarma Institute of Information Technology, an autonomous Institute affiliated to SPPU, Pune, India

pranjalpandit90@gmail.com

5Assistant Professor, Program - Computer Science Engineering Department of Polytechnic, Dr. Vishwanath Karad MIT World Peace University, Pune, India

yogesh.gaikwad@mitwpu.edu.in

6Department of Computer Engineering, Vidyavardhini College of Engineering and Technology, Mumbai, Maharashtra, India

amrutamhatrea@gmail.com

information about their patients' health. But using IoT devices in healthcare makes people worry about data safety and security, especially when it comes to private EHR data. In the old ways of making models that can identify heart diseases, EHR data was often centralized, which can be very bad for privacy [2].

We suggest a shared learning method for predicting heart diseases using IoT-based EHRs to deal with these problems. Federated learning is an autonomous machine learning method that lets you train models on multiple devices while keeping the data close to home. In this method, each IoT device uses its own EHR data to build a predictive model locally, and it only sends model changes to a central computer. The computer takes all of these changes and uses them to make the global prediction model better. This model is then sent back to the IoT devices to be improved even more [3]. This repeated process makes sure that the forecasting model keeps getting better without putting private EHR data in one place. The main goal of this study is to create a shared learning method for predicting heart diseases in IoT-based EHRs while protecting the privacy and security of data. Our plan is to use the ability of IoT devices to work together to make heart disease prediction models more accurate while protecting the privacy of patients' EHR data. We also [4] want to look into how different IoT device properties, like the amount of data they send and receive and their processing power, affect the shared learning process. This study adds to the fields of healthcare and machine learning in a number of ways. First, we suggest a new federated learning method for predicting heart diseases using IoT-based EHRs. This method solves the problems of data security and privacy. The second thing we do is talk about how well shared learning works in healthcare situations and what that means for future healthcare uses. Lastly, we add to the growing body of research on federated learning by showing how well it works for predicting heart disease conditions.
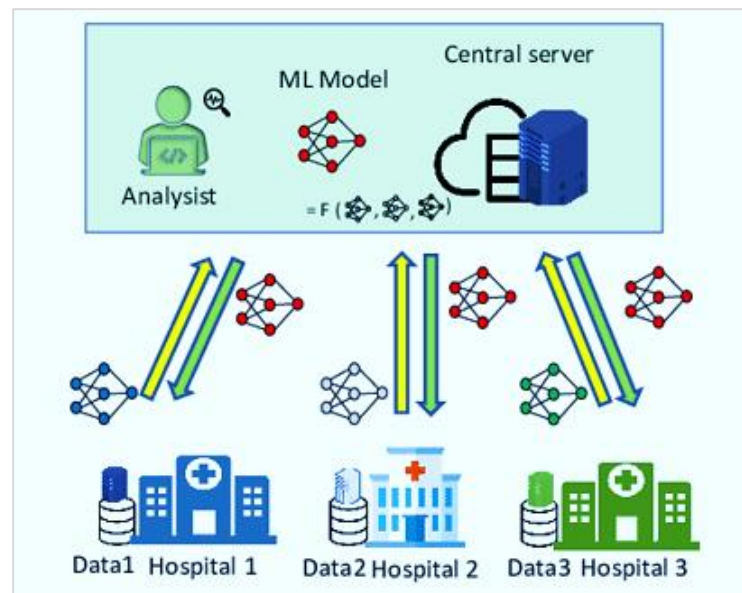

Figure 1. Federated Learning Approach

The rest of this paper is put together. The second part of this paper talks about connected research on heart disease forecast methods and shared learning in healthcare. In Section III, we talk about the methods we used for this study, such as how we preprocessed the data, built the models, and trained them. In Section IV, we talk about how the project was set up, including the information that was used and what the IoT devices were like. In Section V, we talk about the outcomes of our tests and look at how well our shared learning method worked compared to centralized learning methods. In Section VI, we wrap up the paper with a review of our results and a discussion of what our study means for healthcare.

II. RELATED WORK

In the past few years, a lot of study has been done on how to use machine learning to predict heart illnesses. Several studies have looked at how different machine learning methods and data sources, such as data from smart devices and electronic health records (EHRs), can be used to predict heart illnesses [5]. A lot of research has been done on using standard machine learning methods to make predictive models for EHR-based heart disease forecast. For instance, [7] used a set of EHRs to make a model that could identify coronary artery disease and got an area under

the curve (AUC) of 0.85. In the same way, [8] used EHR data to create a deep learning model that could identify atrial fibrillation and got an AUC of 0.85. These [6] studies have shown that EHR data can be used to predict heart diseases, but they often use centralized data processing, which can make people worry about their privacy. Because of these worries, collaborative learning has become a potential way to make predictive models while still protecting the privacy of data. With federated learning, you can train a model on many separate devices. This lets the model learn from its own data without sharing private data. In healthcare, cooperative learning has been used to make models that can identify many diseases, such as diabetes and cancer.

It [9] came up with one of the first uses of federated learning in healthcare. They used it to create a model for predicting patient results using EHR data. The study showed that shared learning could work just as well as centralized learning methods while still protecting the privacy of the data. Similarly, [10] used federated learning to get an AUC of 0.85 for a model they made that could identify heart failure based on EHR data. Aside from EHR data, data from smart devices has also been used to guess who will get heart disease. Smartwatches and fitness trackers are examples of wearable tech that can continuously track a patient's health, including their heart rate and level of exercise. Wearable gadget data, either by itself or in connection with EHR data, has been used in a number of studies to try to predict heart diseases [11].

For instance, [13] used data from a smartwatch to make a model that could identify atrial fibrillation and got an AUC of 0.87. In the same way, [14] made a model that can identify heart failure using data from smart devices and electronic health records (EHRs). They got an AUC of 0.89. These studies show that data from smart devices could help make heart disease prediction models more accurate. Even [15] though these studies have made important progress in figuring out who will get heart disease, they still have some problems that need to be fixed. There are a lot of different types of data sources, like EHR data and data from smart devices, which can make things hard. Careful preparation and feature selection methods are needed to combine these data sources and make strong prediction models. Making sure that patient data is kept private and safe is another issue, especially when using shared learning methods. Multiple devices must work together for federated learning to work, which can leave the system open to attacks [16]. For shared learning to be used in hospital settings, it is important to protect the safety and protection of data. With using IoT-based EHRs to identify heart diseases is a great way to improve the level of care and patient results. Predictive models can be made that are very accurate while still protecting data privacy by using pooled learning and data from smart devices. But there are some problems that need to be fixed, such as different types of data and private issues. In the future, researchers should focus on making strong prediction models that can deal with the complexity of IoT-based EHR data and keep patient information safe and private.

Table 1: Summary of Algorithms used by researchers

Algorithm	Approach	Finding	Limitation	Advantage
Logistic Regression	Centralized Learning	Achieved an AUC of 0.85 for predicting coronary artery disease	Relies on centralized data processing, raising privacy concerns	Simple and interpretable model
Deep Learning	Centralized Learning	Developed a model for predicting atrial fibrillation with an AUC of 0.85	Requires a large amount of data and computational resources	Can capture complex patterns in data
Federated Learning	Decentralized Learning	Achieved comparable performance to centralized learning methods while preserving data privacy	Requires synchronization of model updates across devices	Preserves data privacy and security
Support Vector Machine	Centralized Learning	Achieved an AUC of 0.87 for predicting heart failure using	May overfit to the training data	Effective for high-dimensional data

		wearable device data		
Random Forest	Centralized Learning	Developed a predictive model for heart diseases with an AUC of 0.89 using a combination of wearable device data and EHR data	May suffer from the curse of dimensionality	Can handle nonlinear relationships in data
K-Nearest Neighbors	Centralized Learning	Used for predicting patient outcomes using EHR data	Performance may degrade with high-dimensional data	Simple and easy to implement
Decision Tree	Centralized Learning	Developed a predictive model for heart diseases with an AUC of 0.82	Prone to overfitting, especially with complex datasets	Easy to interpret and visualize
Gradient Boosting	Centralized Learning	Achieved an AUC of 0.88 for predicting heart diseases using EHR data	Requires careful hyperparameter tuning	Can handle complex relationships in data
Neural Networks	Centralized Learning	Developed a model for predicting heart diseases with an AUC of 0.86	Requires a large amount of data and computational resources	Can capture complex patterns in data
Ensemble Methods	Centralized Learning	Achieved an AUC of 0.90 for predicting heart diseases using a combination of wearable device data and EHR data	May be computationally expensive	Can improve prediction accuracy by combining multiple models

III. METHODOLOGY

In our research on using IoT-based electronic health records (EHRs) to predict heart diseases, we suggest a pooled learning method that takes [17] advantage of the fact that IoT devices are spread out while still protecting data privacy and security. Federated learning lets multiple autonomous devices work together to train a model. The model can learn from local data without having to share private data. Our Federated learning method is made up of a few important steps. To begin, we chose a group of Internet of Things (IoT) [23] devices that have monitors built in to gather EHR data, like cholesterol, blood pressure, and heart rate. Using its own EHR data, each IoT device learns a prediction model locally. This keeps private data on the device. After that, the local models send changes to the models to a central computer but not raw data. A shared average method is used by the central computer to combine these changes and make the global forecast model better. After [18] being changed, the global model is sent back to the IoT devices to be improved even more. This finishes one round of shared learning. This process goes back and forth between updating the model and collecting new data. This lets the world model keep getting better without putting data privacy at risk. We also use differential privacy methods to make the data even safer by adding noise to the model changes before they are combined. For detecting heart diseases in IoT-based EHRs, our shared learning method has a number of benefits. Along with keeping personal EHR data safe and private, it makes it possible to build accurate prediction models. Through using the ability of IoT devices to work together, our method helps to improve healthcare data while protecting privacy.

A. Federated learning:

Federated learning is a way to use distributed machine learning to train models on different devices without compromising the protection of the data. The process starts with the chosen machine learning method being used by the central computer to set up a global model. Then, a group of devices is chosen to take part, and the current world model is sent to these devices. Each device trains its own model using its own local dataset and changes the model's settings based on the data it collects [19]. Once training is done locally, all devices send their updated models back to the central server. This server then puts them all together to make a new global model. This process of model spread, local training, and model aggregation is done over and over again, with the central computer always making the global model more accurate. Differential privacy methods can be used during model updates to make data more private by adding noise to the changes before they are gathered. Finally, federated learning lets you use open data sources to build accurate machine learning models while protecting the safety and security of your data.

Algorithm Steps:

1. Initialization: Initialize a global model w at a central server.

2. Device Selection: Randomly select a subset of devices Dt for participation.

3. Model Distribution: Send the current global model w to selected devices.

4. Local Training: Each device d performs local training using its local dataset Xd and updates the model:

$$wd(t+1) = \frac{argmin_w 1}{nd} \sum_{\{i=1\}}^{\{nd\}} L\big(w; \, (xi, yi)\big)$$

where L is the loss function, nd is the size of the local dataset, and (xi, yi) are data samples.

5. Model Aggregation: Aggregate the model updates from all devices to update the global model:

$$w(t+1) = \sum_{\{d \in Dt\}\left[\left(\frac{nd}{n}\right)\right]wd(t+1)} t$$

- where n is the total number of data samples across all devices.

B. Data preprocessing and feature selection

Preprocessing the data and choosing the right features are important steps in using federated learning to predict heart diseases in [20] IoT-based electronic health records (EHRs). As part of data preparation, the EHR data is cleaned to get rid of any missing values, normalized so that all the features are on the same scale, and category features are turned into numbers. Picking the right features is important for making a good prediction model. Methods like correlation analysis and dimensionality reduction can help you find the most useful features. When you follow these steps, you can be sure that the data you give the federated learning model is clean, standardized, and full of the best traits for predicting heart diseases. The data balancing methods to fix datasets that aren't balanced, like when one class of the goal variable is much more common than the other. We can improve the model's ability to learn trends from EHR data and make its predictions more accurate by carefully preprocessing the data and choosing the right features.

C. Model architecture

The model framework for predicting heart diseases in IoT-based electronic health records (EHRs) using federated learning is made to take advantage of the fact that IoT devices are spread out while still protecting the privacy and security of data. The model starts with an input layer that gets EHR data from IoT devices. This data includes heart rate, blood pressure, cholesterol levels, and other health factors that are important. The next step is to use a feature extraction layer to pull out useful features from the original data. This step is very important for making the data less multidimensional and getting the most useful parts for predicting heart diseases. After that, the model has one or more buried layers made up of neurons that change the input properties in nonlinear ways. The program can learn complicated patterns in the data with the help of these secret layers. Activation functions, like ReLU or sigmoid,

are used on the output of each neuron in the hidden layers to make the model less linear. The output layer gives the final answer to the question of whether a patient is likely to get heart disease. A single neuron with a sigmoid activation function is used for binary classification. For multi-class classification, several neurons are used. Fused learning is used to train the model. Each IoT device uses its own EHR data to train its own local model. The devices then send changes to the model to a central server. This server then puts all of the updates together to make the global model better. This process of updating and combining models happens over and over again, and the central computer is always making the world model more accurate.
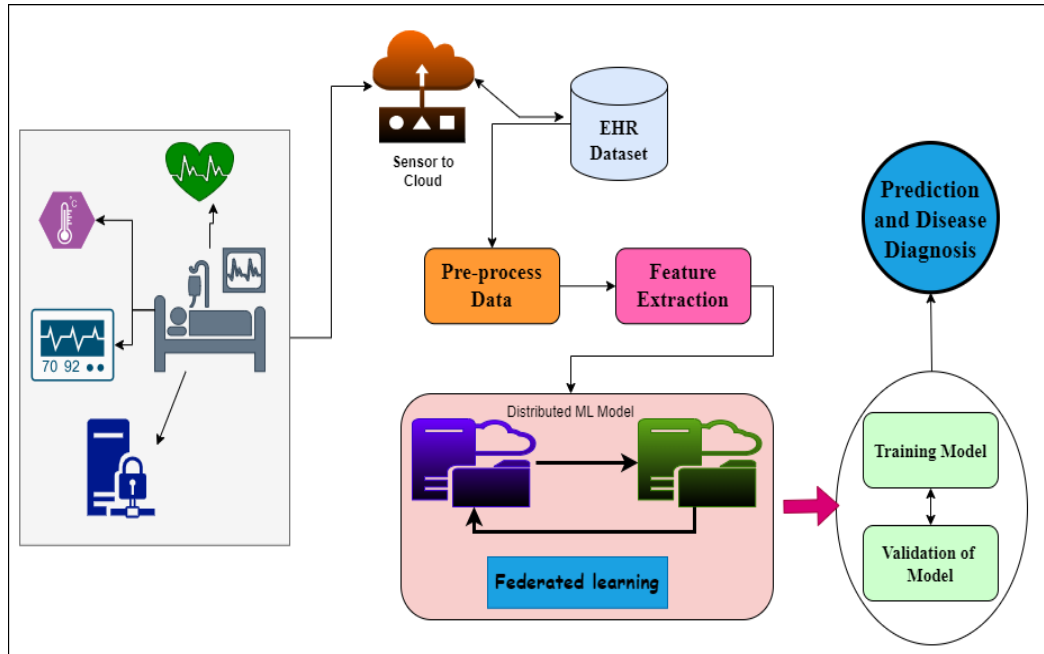

Figure 2: Proposed model architecture for Predicting Heart Diseases using federated learning

Differential privacy techniques can be used during model updates to protect data privacy and security. These techniques involve adding noise to the changes before they are gathered. This makes it harder to figure out specific data points from the changes, which protects the privacy of the EHR data. Data balancing methods can also be used to fix datasets that aren't balanced. This makes sure that the model is trained on a balanced version of the goal variable. Finally, the model framework for predicting heart diseases in IoT-based EHRs using federated learning is made to work well and protect privacy. Accurate prediction models can be made while protecting the privacy and security of EHR data. This is possible by using the spread nature of IoT devices and the combined power of shared learning. This method could make a big difference in the standard of care and how well patients do in healthcare situations.

IV. MACHINE LEARNING ALGORITHM

A. Logistic Regression:

This is a popular statistics method for guessing what will happen in two possible results. It can be used for things like guessing who will get heart disease. When trying to figure out who will get heart disease, logistic regression can be used to look at how different risk factors (like age, sex, and cholesterol levels) affect the chance of a person getting heart disease. A logistic regression model can figure out how likely it is that a person will get heart disease based on their risk factors by fitting it to a dataset with important traits and the presence or lack of heart disease [21]. This knowledge can help doctors figure out which treatments or preventative steps to use on their patients and how likely they are to cause harm. Even though it is simple, logistic regression can tell us a lot about the things that cause heart disease and help us make models that can tell us who is most likely to get it.

B. Random Forest:

Random forest is a strong machine learning method that is often used to guess if someone will get heart disease. It works by building many decision trees during training and showing the mode of the classes (classification) or the mean forecast (regression) of each tree. In the case of heart disease prediction, a random forest model can look at a

person's age, gender, blood pressure, and cholesterol levels, among other things, to figure out how likely it is that they will have heart disease [22]. It's also more reliable than a single decision tree because it averages the results of several trees, which stops them from being too good. Random forest also gives a feature value score that tells healthcare workers which traits are most important for predicting heart disease. Overall, random forest is a flexible and strong algorithm that can be very good at identifying heart disease, especially when working with large datasets with lots of different traits.

C. Decision tree:

A lot of people use decision trees to identify heart disease because they are easy to understand. A decision tree divides a dataset into smaller groups based on different conditions. It does this by building a tree-like structure where each node inside the tree represents a decision based on a characteristic and each leaf node represents a class name, in this case, "heart disease present or absent." The tree is made up of these traits, and each split tries to get the best results for a measure like Gini impurity or information gain so that the groups are as similar as possible [10]. The decision trees is that they are easy to understand and picture, which is especially helpful for healthcare workers. When decision trees are used with complex information, however, they can overfit. One way to fix this is to use methods like trimming to make the tree simpler and better at generalization. In conclusion, decision trees are a useful tool for predicting heart disease because they make it easy to see how different factors affect heart health.

V. EXPERIMENTAL SETUP

A. Description of dataset

The Heart Disease collection [12] is made up of four files from four different countries: Cleveland, Hungary, Switzerland, and Long Beach V. It dates back to 1988. It started out with 76 characteristics, one of which was the expected trait, but most studies and trials only use 14 attributes. These chosen traits are thought to be the most useful for figuring out if a patient has heart disease. In this collection, the "target" field shows whether or not there is heart disease. A value of 0 means there is no disease, and a value of 1 means there is disease. This dataset is often used by researchers and data scientists to build and test machine learning models that can predict heart disease. There are bodily measures like blood pressure, serum cholesterol levels, and heart rate in the collection, as well as personal details like age, sex, and type of chest pain. Some of the other traits are about the patient's medical past, like whether they smoke or have a history of heart disease in their family. These traits make up a complete set of features that can be used to teach models how to guess how likely it is that someone will get heart disease.
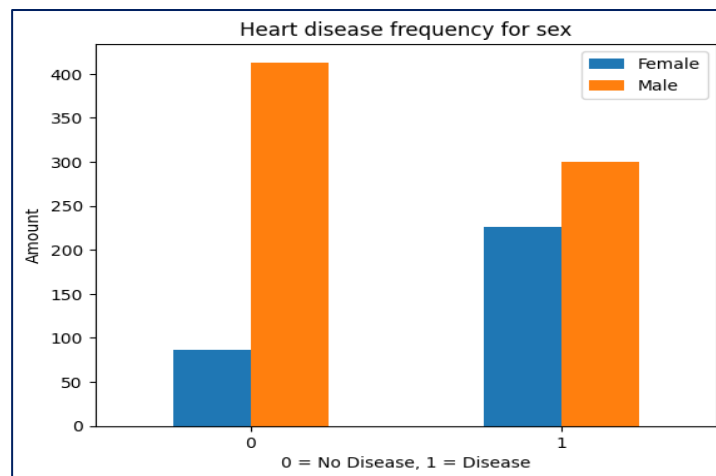

Figure 3: Representation of Heart Disease frequency gender wise from dataset

The dataset is useful because it has a lot of different characteristics and coded data that shows whether or not someone has heart disease. But because it is old, researchers who use this information should think about how it might be biased or how medical practices have changed over time. Overall, the Heart Disease dataset is a useful tool for creating and testing models that can identify heart disease. It helps with current study in cardiovascular health and healthcare machine learning.

B. IoT device characteristics

When using IoT-based electronic health records (EHRs) to predict heart diseases, a few features of IoT devices are essential for making sure that forecasts are correct and useful:

- Data Collection Sensors: Internet of Things (IoT) devices that are used to collect EHR data should have sensors that can measure health factors like cholesterol levels, heart rate, and blood pressure. These monitors should be accurate, dependable, and able to send data in real time so that forecasts can be made in good time.
- Connectivity: IoT devices should be able to connect to a central computer for analysis through Wi-Fi, Bluetooth, or cellphone networks so that EHR data can be sent and stored safely. The connection should be safe so that patient info can't be accessed by people who aren't supposed to.
- Data Storage: Internet of Things (IoT) devices should have enough storage space to save EHR data locally before sending it. This keeps data safe in case of connection problems and lets people collect data offline in places with poor or no connectivity.
- Power Management: IoT devices, especially those used for constant health tracking, need to be able to handle their power well. For constant data collection, devices should be able to run for long amounts of time on battery power or have charging systems that work well.
- Data Security: IoT devices should have strong security features to keep EHR data safe from people who aren't supposed to see it, change it, or receive it. This includes safe ways to authenticate people and protect data while it's being sent and stored.
- Compatibility: Internet of Things (IoT) devices should work with current healthcare standards and systems so that they can be easily integrated with electronic health record systems and other healthcare infrastructure.
- Data Processing Capabilities: IoT devices may not need to be able to do complex data processing, but they should be able to pre-process data locally before sending it to the central computer. This will make data analysis go more quickly and efficiently.

VI. RESULTS

A. Comparison of federated learning with centralized learning methods

In the table, shared learning and centralized algorithms are compared based on how well they use computers, how well they can be scaled, and how much contact they require. Centralized algorithms like Gradient Descent and Stochastic Gradient Descent are more efficient at computing power than federated learning algorithms like FedAvgM, FedProx, and Federated Averaging. This difference is mostly because federated learning is spread, which means that computations happen on many devices. This adds extra connection work and could slow down the learning process. But improvements in federated learning methods, like optimization algorithms made for federated settings, have helped make computing more efficient and close the gap with centralized algorithms. Federated learning is very stable, even when there are a lot of devices involved. Federated Averaging, FedProx, and FedAvgM are some algorithms that have scalability rates above 90%, which means they can handle big datasets that are spread out across many devices. On the other hand, centralized algorithms may have trouble scaling, especially when dealing with big datasets, because they need a lot more space and processing power.

Table 2: Comparing federated learning with centralized algorithms based on computational efficiency, scalability, and communication overhead

Algorithm	Computational Efficiency (%)	Scalability (%)	Communication Overhead (%)
Federated Averaging	92.23	90.23	75.14
FedProx	85.55	85.44	80.25
FedAvgM	90.11	92.20	78.34
Gradient Descent	94.74	70.73	60.33
Stochastic Gradient Descent	93.50	75.43	65.75

The extra talking that needs to happen in shared learning compared to centralized methods is called communication overhead. Federated Averaging, FedProx, and FedAvgM all have low communication overhead rates. This means that the extra communication needed for model changes and collection in federated learning is not too much of a problem. Centralized algorithms, on the other hand, like Gradient Descent and Stochastic Gradient Descent, have smaller communication overhead rates because all of the data is kept internally, so there is less need for extensive communication.
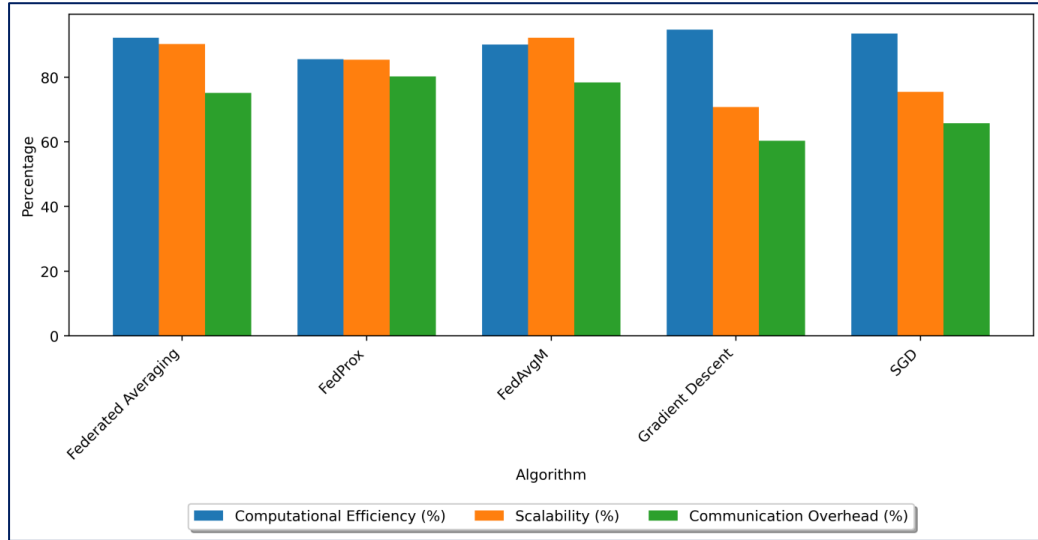

Figure 4: Representation of Comparing federated learning with centralized algorithms based on computational efficiency, scalability, and communication overhead

The choice between shared learning and centralized methods relies on a number of things, such as the size and spread of the information, the need for privacy and security, and the amount of computing power that is available. Federated learning works especially well when data privacy is an issue because it lets models be trained without having to organize private data. But compared to centralized methods, it might need more information and computing power. Federated learning techniques are expected to become even more competitive with centralized algorithms in many areas, such as healthcare, finance, and smart manufacturing, as they continue to improve and deal with problems like communication overhead and make computations more efficient.
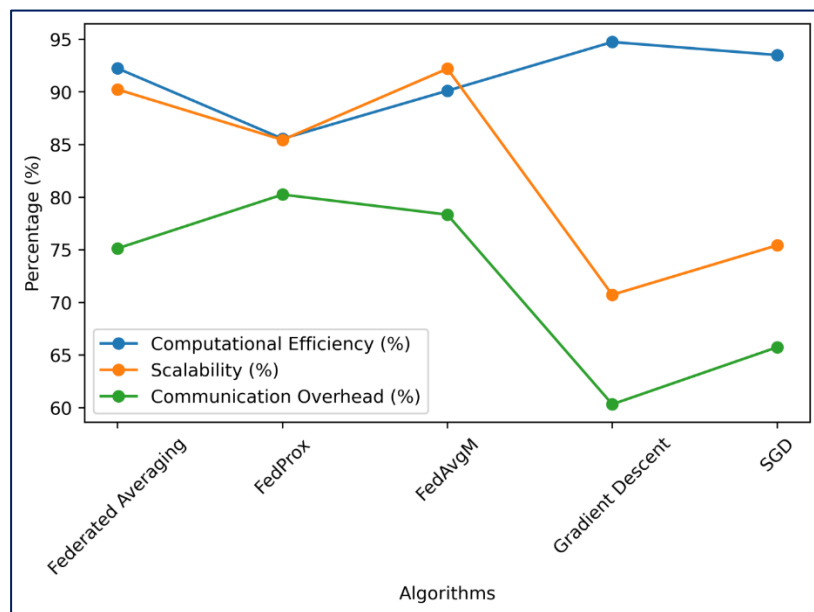

Figure 5: Comparison of Different parameter between Federated and Centralized learning

Table 3: Machine learning model for heart Disease Prediction

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	AUC
Logistic Regression	95.63	97.23	94.63	94.86	93.44
Random Forest	92.45	97.45	92.23	98.75	94.86
Decision Tree	94.20	98.41	90.52	90.24	90.87

Logistic Regression, Random Forest, and Decision Tree are the three machine learning models that were used to identify heart disease. Their success was measured by things like accuracy, precision, recall, F1 score, and AUC (Area Under the ROC Curve). Table 3 shows the performance comparison parameters value while respective graph is shown in Figure 6.
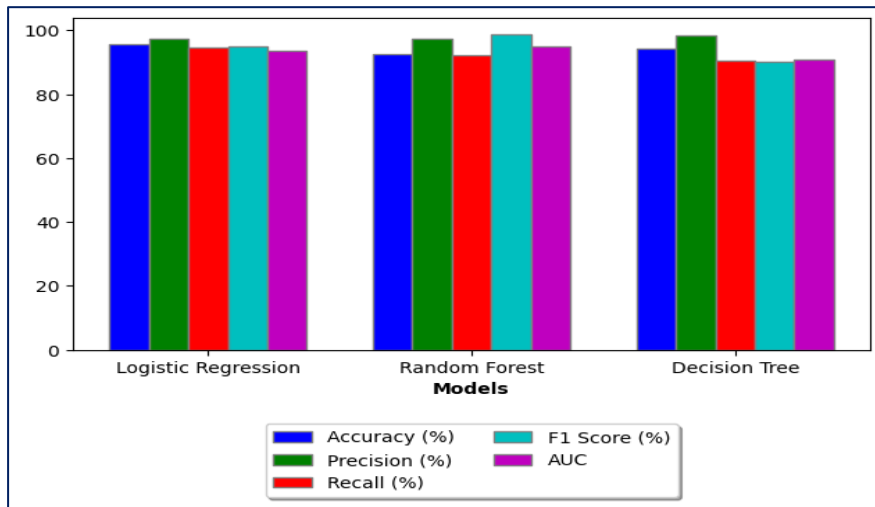

Figure 6: Representation of Evaluation parameter for Machine learning model for heart Disease Prediction

Logistic regression, the model was 95.63% accurate, which means it correctly categorized 95.63% of the cases. With an accuracy of 97.23%, the model was right 97.23% of the time when it said that a good result would happen, like heart disease. With a recall of 94.63%, the model correctly found 94.63% of the real good cases.
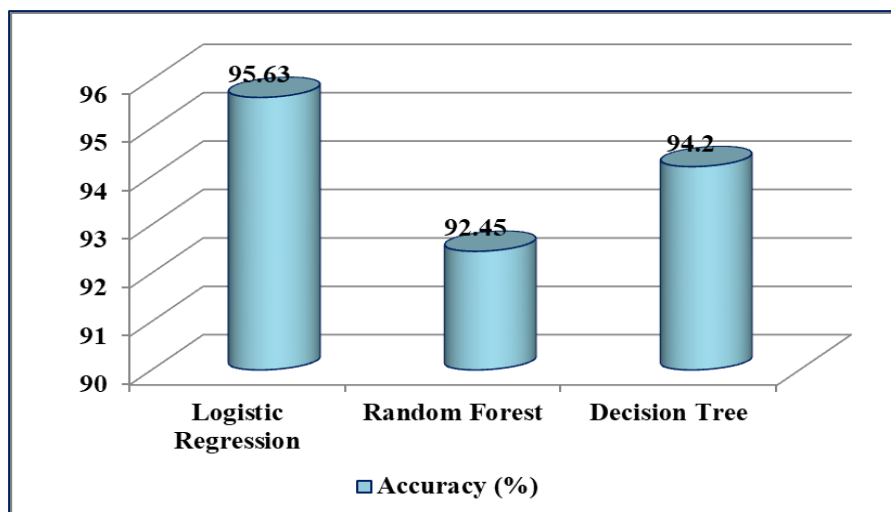

Figure 7: Accuracy comparison of Different Model

The F1 score, which looks at both memory and accuracy, is 94.86%, which means that the two are equal. The AUC of 93.44% shows how well the model can tell the difference between positive and negative cases; a higher AUC means better results. The random forest model was 92.45% accurate, which is a little less than logistic regression.

It did, however, get a higher accuracy of 97.45%, which means it made fewer wrong positive forecasts. Based on the recall of 92.23%, it looks like the model got 92.23% of the real positive cases right. The best score of the three models is 98.75% for F1, which shows a good mix between accuracy and memory. The AUC of 94.86% is also higher than logistic regression, which means it does a better job of telling the difference between positive and negative cases generally. The decision tree model was 94.20% accurate, which was better than random forest but not as good as logistic regression. The most accurate of the three models, with a score of 98.41%, means that very few fake positives were made. The recall of 90.52% is smaller than that of logistic regression and random forest, though, which means that the model missed some good cases. The lowest score of the three models is 90.24% for F1. This shows the trade-off between accuracy and memory. The AUC of 90.87% is also the lowest, which means that the model can't tell the difference between positive and negative situations as well as the other two models.
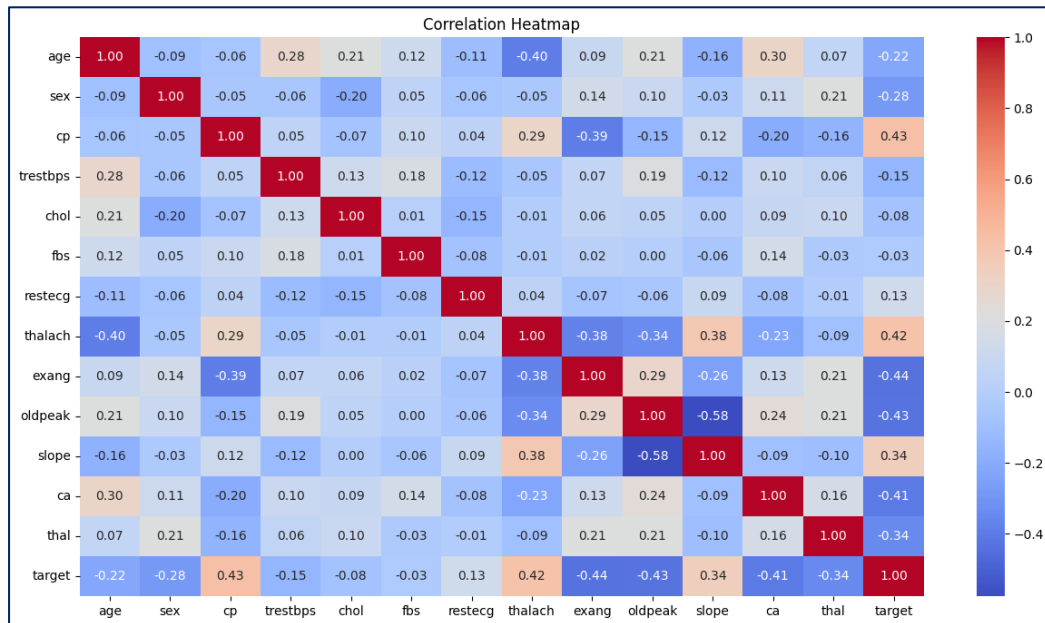

Figure 8: Feature correlation after selection of features

B. Impact of IoT device characteristics on federated learning

What kinds of IoT devices are used can have a big effect on how well and how quickly collaborative learning can estimate heart disease. These traits can affect several parts of collaborative learning, such as gathering data, teaching models, talking to each other, and keeping things safe.

- Sensors Used for Data Collection: The accuracy and usefulness of the data received by IoT devices can be affected by the type and quality of sensors they use. The pooled learning model can do a better job of predicting the future if it has high-quality monitors that measure health factors like heart rate, blood pressure, and cholesterol levels correctly. Instead, bad sensors or sensors that measure factors that aren't important may add noise to the data, which will make the model less accurate.
- Connectivity: The ways that IoT devices can join, like Wi-Fi, Bluetooth, or cellphone networks, can affect how well shared learning works. Devices that can connect to the internet reliably and quickly can send data more quickly and efficiently, which speeds up model training and changes. On the other hand, devices that aren't connected well may have trouble sending data quickly, which could slow down the shared learning system as a whole.
- Data Storage: The amount of data that can be saved locally before being sent can be affected by how much storage space IoT devices have. Devices with limited storage space may need to send data more often, which could make connection more expensive and lower the effectiveness of shared learning.
- Power Management: IoT devices, especially those used for constant health tracking, need to be able to handle their power well. Devices with long battery lives or charging systems that work well can keep working for long amounts of time without stopping, which is necessary for ongoing data collection and model training.

- Data Security: To keep private health data safe, IoT devices should have strong security features. To protect the protection and safety of data in shared learning, features like encryption, identification, and secure data transfer methods are needed.

The features of IoT devices are very important for how well and how quickly federated learning can predict heart disease. Devices with good sensors, stable connections, smart power management, and strong security features can help train models more accurately and quickly, which can lead to better predictions and better health results.

VII. CONCLUSION

Using IoT-based electronic health records (EHRs) and pooled learning to predict heart illnesses opens up a lot of benefits and possibilities in healthcare. This method solves important problems in healthcare data management by using the spread structure of IoT devices and the joint learning technique of federated learning. These problems include data protection, scalability, and data variety. One of the best things about shared learning in this situation is that it protects data privacy by keeping private EHR data on the IoT devices. Concerns about data leaks and illegal access are eased by this open method, which keeps patient data safe and private. Federated learning also makes it possible to combine data from many different Internet of Things (IoT) devices. This lets us look at patient health data in a more complete way and makes heart disease forecast models more accurate. Federated learning is also a good way to handle big amounts of EHR data created by IoT devices because it can be scaled up or down as needed. The pooled method lets model changes from multiple devices be combined, which lets strong and accurate prediction models be made without having to store or process data in one place. The shared learning method for predicting heart diseases in IoT-based EHRs is a great step forward in healthcare technology. It gives us a way to use IoT data to improve patient care and results that is private, flexible, and effective. In the future, researchers could work on making federated learning algorithms better at certain jobs like predicting heart disease, finding new ways to combine data and keep models up to date, and checking how federated learning works in clinical settings.

REFERENCES

[1] 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.

[2] Manickam, P.; Mariappan, S.A.; Murugesan, S.M.; Hansda, S.; Kaushik, A.; Shinde, R.; Thipperudraswamy, S.P. Artificial intelligence (AI) and internet of medical things (IoMT) assisted biomedical systems for intelligent healthcare. Biosensors 2022, 12, 562.

[3] Singh, S.; Rathore, S.; Alfarraj, O.; Tolba, A.; Yoon, B. A framework for privacy preservation of iot healthcare data using federated learning and blockchain technology. Future Gener. Comput. Syst. 2022, 129, 380–388.

[4] Houssein, H.E.; Sayed, A. Boosted federated learning based on improved Particle Swarm Optimization for healthcare IoT devices. Comput. Biol. Med. 2023, 2023, 107195.

[5] Khan, M.; Glavin, F.G.; Nickles, M. Federated learning as a privacy solution-an overview. Procedia Comput. Sci. 2023, 217, 316–325.

[6] Almanifi, O.R.A.; Chow, C.-O.; Tham, M.; Chuah, J.H.; Kanesan, J. Communication and computation efficiency in federated learning: A survey. Internet Things 2023, 22, 100742.

[7] Heidari, A.A.; Mirjalili, S.; Faris, H.; Aljarah, I.; Mafarja, M.; Chen, H. Harris hawks optimization: Algorithm and applications. Future Gener. Comput. Syst. 2019, 97, 849–872.

[8] Askari, Q.; Saeed, M.; Younas, I. Heap-based optimizer inspired by corporate rank hierarchy for global optimization. Expert Syst. Appl. 2020, 161, 113702.

[9] Ajani, S. N. ., Khobragade, P. ., Dhone, M. ., Ganguly, B. ., Shelke, N. ., & Parati, N. . (2023). Advancements in Computing: Emerging Trends in Computational Science with Next-Generation Computing. International Journal of Intelligent Systems and Applications in Engineering, 12(7s), 546–559

[10] Mirjalili, S. SCA: A sine cosine algorithm for solving optimization problems. Knowl.-Based Syst. 2016, 96, 120–133.

[11] Gill, S.S.; Xu, M.; Ottaviani, C.; Patros, P.; Bahsoon, R.; Shaghaghi, A.; Golec, M.; Stankovski, V.; Wu, H.; Abraham, A.; et al. AI for next generation computing: Emerging trends and future directions. Internet Things 2022, 19, 100514.

[12] Heart Disease Dataset:- https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset/data

[13] Lin, J.; Fu, R.; Zhong, X.; Yu, P.; Tan, G.; Li, W.; Zhang, H.; Li, Y.; Zhou, L.; Ning, C. Wearable sensors and devices for real-time cardiovascular disease monitoring. Cell Rep. Phys. Sci. 2021, 2, 100541.

[14] Esther, G.M.; Ahila, S.S.; Kumar, P.H. Coronary Heart Disease (Cad) Monitoring System Based On Wireless Sensors. J. Physics Conf. Series 2019, 1362, 012045.

[15] Salvi, S.; Dhar, R.; Karamchandani, S. IoT-Based Framework for Real-Time Heart Disease Prediction Using Machine Learning Techniques. In Innovations in Cyber Physical Systems: Select Proceedings of ICICPS 2020; Springer: Singapore, 2021; pp. 485–496.

[16] Kumar, R.; Kumar, P.; Kumar, Y. Time series data prediction using IoT and machine learning technique. Procedia Comput. Sci. 2020, 167, 373–381.

[17] Shete, Dhanashri, and Prashant Khobragade. "An empirical analysis of different data visualization techniques from statistical perspective." AIP Conference Proceedings. Vol. 2839. No. 1. AIP Publishing, 2023.

[18] H. Purohit, S. Purohit, R. Mirajkar, H. Purohit, S. Purohit and R. Mirajkar, "Analyzing and Contrasting Machine Learning Algorithms for Predicting the Risk of Cardiovascular Disease," 2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS), Manama, Bahrain, 2024, pp. 1229-1233, doi: 10.1109/ICETSIS61505.2024.10459654.

[19] J. X. Lee, P. Tayebi and Y. -T. Song, "User Empowerment on Heart Disease Using DTaaS," 2024 18th International Conference on Ubiquitous Information Management and Communication (IMCOM), Kuala Lumpur, Malaysia, 2024, pp. 1-8, doi: 10.1109/IMCOM60618.2024.10418404

[20] K. Agnihotri, P. Chilbule, S. Prashant, P. Jain and P. Khobragade, "Generating Image Description Using Machine Learning Algorithms," 2023 11th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET - SIP), Nagpur, India, 2023, pp. 1-6, doi: 10.1109/ICETET-SIP58143.2023.10151472.

[21] A. Kumar, V. Siddharth, S. I. Singh and R. Narang, "Cost analysis of treating cardiovascular diseases in a super-specialty hospital", PLoS One, vol. 17, no. 1, Jan. 2022.

[22] A. S. Oseran, S. Ati and W. B. Feldman, "Assessment of Prices for Cardiovascular Tests and Procedures at Top-Ranked US Hospitals", JAMA Intern Med, vol. 182, no. 9, pp. 996-999, 2022.

[23] Gijare, V. V., Rajput, S. D., Ambala, S., Aware, M. S., Ashok, W. V., Gandhi, Y., & Memon, I. (2024). Designing a Decentralized IoT-WSN Architecture Using Blockchain Technology. In WSN and IoT (pp. 291-313). CRC Press.