Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi
Afyon Kocatepe University – Journal of Science and Engineering
https://dergipark.org.tr/tr/pub/akufemubid

Araştırma Makalesi / Research Article
e-ISSN: 2149-3367
DOI: https://doi.org/10.35414/akufemubid.1679417
CITE: Afyon Kocatepe University Journal of Science and Engineering 26(2) (2026) 337-353

Decentralized Disease Prediction: A Federated Learning Perspective

*Makale Bilgisi / Article Info
Alındı/Received: 18.04.2025
Kabul/Accepted: 09.01.2026
Yayımlandı/Published: 08.04.2026

Merkeziyetsiz Hastalık Tahmini: Federe Öğrenme Yaklaşımı

Abdulbaki DEMIR[1], Ahmet Yasir KULAKSIZ[1], Busra BUYUKTANIR[1], Gozde Karatas BAYDOGMUS[2], Kazim YILDIZ[1*]

[1] *Marmara University, Faculty of Technology, Department of Computer Engineering, Istanbul, Turkey*
[2] *Loyola University, Computer Science Department, Chicago, USA*

Abstract

Machine learning is crucial in healthcare for analyzing large datasets. However, privacy concerns regarding sensitive health data limit centralized collection and standard ML usage. To enhance privacy, this study uses the Data Unaware Classification Based on Association (duCBA) algorithm to aggregate classification rules rather than model weights. This approach ensures data security and offers interpretable disease predictions. Model training was conducted using two publicly available health datasets: Diabetes Health Indicators and Heart Attack Prediction. Prior to training, feature selection techniques including Mutual Information, Chi-Square, and Principal Component Analysis (PCA) were applied to generate three distinct datasets. The system architecture comprises a central server, an API server, and multiple clients. Clients locally train models on their own data without sharing it externally. The API server collects these locally trained models and transmits them to the central server. The central server then aggregates them using the duCBA algorithm and disseminates the updated global model back to the clients. This research not only demonstrates the practical viability of duCBA in federated settings but also compares its performance against the widely-used Federated Averaging (FedAVG) algorithm. Experimental results reveal that duCBA achieves comparable performance to FedAVG while offering improved model transparency in line with Explainable Artificial Intelligence (XAI) principles. Notably, the duCBA algorithm delivers stable and consistent across different feature selection methods and varying client numbers, particularly excelling in scenarios with fewer clients.

Keywords: *Federated Learning; Relational Classification; duCBA; Data Privacy; Health Datasets*

Öz

Makine öğrenimi, büyük verilerden anlamlı bilgi çıkarabildiği için sağlık sektöründe hayati öneme sahiptir. Ancak, hassas sağlık verilerine dair gizlilik endişeleri, merkezi veri toplamayı ve standart algoritmaların kullanımını zorlaştırmaktadır. Bu çalışma, gizliliği sağlamak için model ağırlıkları yerine sınıflandırma kurallarını birleştiren Data Unaware Classification Based on Association (duCBA) algoritmasını kullanır. Bu yöntem veri güvenliğini artırır ve hastalık tahmininde açıklanabilir sonuçlar sunar. Model eğitimi, Diabetes Health Indicators ve Heart Attack Prediction adlı iki açık erişimli sağlık veriseti kullanılarak gerçekleştirilmiştir. Eğitime başlamadan önce Mutual Information, Chi-Square ve Principal Component Analysis (PCA) gibi öznitelik seçimi yöntemleri uygulanarak üç farklı veri kümesi oluşturulmuştur. Sistem mimarisi; bir merkezî sunucu, bir API sunucusu ve birden fazla istemciden oluşmaktadır. İstemciler, verilerini paylaşmadan kendi cihazlarında yerel modelleri eğitmektedir. API sunucusu, bu yerel modelleri toplayarak merkezî sunucuya iletmektedir. Merkezî sunucu ise bu modelleri duCBA algoritması ile birleştirerek güncel küresel modeli istemcilere tekrar dağıtmaktadır. Bu çalışma, duCBA algoritmasının federe ortamlarda uygulanabilirliğini ortaya koymakla kalmayıp, yaygın olarak kullanılan Federated Averaging (FedAVG) algoritması ile performans karşılaştırmasını da içermektedir. Sonuçlar, duCBA'nın FedAVG ile benzer performans sergilediğini ve aynı zamanda Açıklanabilir Yapay Zekâ (Explainable Artificial Intelligence, XAI) ilkeleriyle uyumlu olarak modelin karar süreçlerini şeffaf biçimde izlemeye olanak tanıdığını göstermektedir. Farklı öznitelik seçimi yöntemleri ve istemci sayılarıyla yapılan testlerde, özellikle az sayıda istemcinin bulunduğu sistemlerde duCBA algoritmasının istikrarlı ve dengeli bir başarı sağladığı gözlemlenmiştir.

Anahtar Kelimeler: *Federe Öğrenme; İlişkisel Sınıflama; duCBA; Veri Mahremiyeti; Sağlık Verileri*

1. Introduction

The With the rapid development of data-driven technologies, machine learning models play an important role in many fields (Sasaki 2022). These data are evaluated in many different areas, from analyzing user movements to monitoring environmental factors (John Dian et al. 2020). In addition, various systems are developed thanks to the predictive models created with machine learning algorithms (Bian et al. 2022). By analyzing large amounts of data, these models can reveal complex patterns and relationships, predict future outcomes and provide critical information for decision support systems. However, large and diverse datasets are needed for machine learning models to work effectively

*Sorumlu Yazar/Corresponding Author: Kazım YILDIZ
e-posta/e-mail: kazim.yildiz@marmara.edu.tr

(Surur et al. 2025). This situation brings along problems such as hardware capacity strain, prolonged training time and data security with the increase in data size (Paleyes et al. 2022).

Data privacy and security are of critical importance, especially in areas with sensitive data such as the healthcare industry (Choudhury et al. 2019, Zhou and Li 2023). The collection of health data in a centralized server faces limitations due to both legal regulations and security risks (Sun et al. 2023). For example, personal data protection regulations in many countries restrict the sharing and centralized processing of health data. This limits the applicability of traditional machine learning methods and necessitates the development of new approaches.

In this context, federated learning stands out as an innovative solution that eliminates the need for centralized data collection and enables data to be processed on local devices. Federated learning allows machine learning models to be trained in a distributed manner while maintaining the privacy of user data (Raj et al. 2022, Buyuktanir et al. 2025). This method offers a powerful alternative, especially in areas where data privacy is at the forefront, such as the healthcare industry. Federated learning in healthcare holds significant potential by enabling large-scale, multi-center research and collaborations without compromising data security and privacy (Li et al. 2025). With federated learning, patient data is processed on local devices and only model updates are transmitted to the central server. Thus, communication and storage costs are reduced (Li et al. 2020). Moreover, since the federated learning method does not send data to the centralized structure, the security of critical information on local devices is also ensured (Wen et al. 2023).

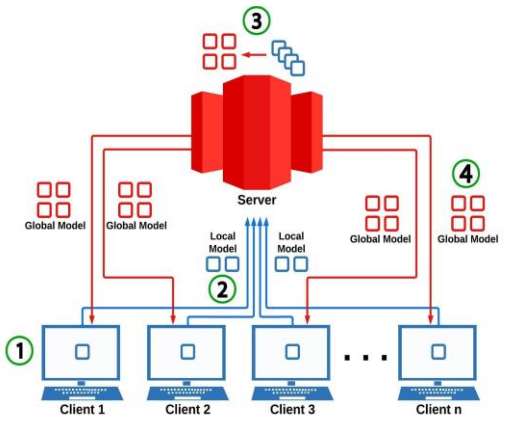

Figure 1. Federated Learning Workflow Diagram

Another important advantage of federated learning is that it allows for more generalizable models by increasing data diversity. While in traditional methods, aggregating data from different sources creates legal and technical challenges, federated learning overcomes these obstacles by enabling local processing of this data. It also increases trust and encourages wider participation by ensuring that data owners retain control.

Figure 1 shows the workflow of the federated learning architecture. The flow works as follows:

- Each client trains a local model with its own data.
- The trained local models are sent to a central server. However, only the updated parameters of the models are shared, not the data.
- The central server builds a global model by merging the local models received from all clients using federated aggregation algorithms (Onlu et al. 2025). Through this process, the knowledge learned by each client is effectively merged into a single, unified global model.
- The server resends the created global model to all clients. Clients receive this model and retrain it with their local data in the next round.

Federated learning architectures are used in many areas from education to health, security and finance (Çıplak et al. 2025, Ozturk et al. 2025, Büyüktanir et al. 2025). Especially in terms of protecting patient privacy, researches in the field of health should be important and recent. In this research, an application on disease prediction in health data using a federated learning architecture is developed. Medical institutions can collaboratively develop AI models through Federated Learning by utilizing data protection techniques; this allows them to support rapid clinical decision-making processes while ensuring patient privacy (Albshaier et al. 2025). Diabetes is an increasing health problem worldwide and its early diagnosis is of great importance (Ali et al. 2022). Additionally American Heart Association (AHA) Presidential Advisory warns that cardiovascular disease is projected to rise from 11.3% to 15% by 2050, affecting 45 million U.S. adults (American Heart Association, 2024). However, the privacy and security of medical data is a major obstacle to the application of traditional machine learning approaches. With federated learning, patient data is processed on local devices without transferring it to a central server, thus protecting patient privacy.

1. 1 Main Contributions

The main contributions of this study can be summarized as follows:

- Proposing an Innovative Integration with Federated Learning: In this study, the Data-Unaware Classification Based on Association (duCBA) algorithm (Büyüktanır et

al. 2023) one of the federated aggregation methods was employed. duCBA is grounded in relational classification and generates labeled association rules (Tosun et al. 2025). Owing to this structure, the global model constructed within the federated setting not only preserves data privacy but also provides an interpretable and explainable decision-making process through its rule-based framework.

- Application and Validation in the Healthcare Domain: The research demonstrates the viability of the proposed framework on two critical health datasets: the Diabetes Health Indicators Dataset (Centers for Disease Control and Prevention 2024) and the Heart Attack Dataset (Fayez 2024). The consistent results across both datasets underline the adaptability and robustness of the duCBA algorithm for risk analysis and early disease detection.
- Enhancing Predictive Performance through Advanced Feature Selection: Leveraging advanced feature selection methods, the study refines model inputs to achieve superior predictive accuracy. This ensures the reliability of disease predictions while preserving data privacy, a critical factor in healthcare applications.
- Providing a Scalable and Privacy-Preserving Solution: By addressing challenges such as data security and federated model scalability, the study offers a practical solution for deploying federated learning in sensitive domains like healthcare, paving the way for further real-world applications (Kucur et al. 2026).

The next sections of the research examine the following topics. Section 2 reviews the work that has been done for blended learning. Section 3 details the implementation techniques developed. Section 4 presents the experimental results from the implementation. Section 5 discusses the research findings and Section 6 summarizes the overall conclusions.

2. Related Works

The federated learning architecture developed by Google in 2016 to protect data privacy offers various advantages. Thanks to this architecture, data privacy is ensured during model training, while training processes are accelerated. Instead of data, sending the model between the clients and the server, rather than the data, avoids challenges like bandwidth limitations and communication latency. This approach also reduces the need for data storage on the central server, thereby lowering costs. However, besides these advantages of federated learning, there are also important problems that are discussed under four main headings: communication overhead, system heterogeneity, statistical heterogeneity and privacy concerns.

Managing the process of submitting models becomes increasingly challenging as the number of local devices grows, which is referred to as communication overhead (Li et al. 2020). System heterogeneity refers to variations in performance caused by differences in the hardware and software of the devices used for model training (Wen et al. 2023). Statistical heterogeneity arises from differences in the data distributions collected by the devices, which can lead to the model performing well on some devices but poorly on others (Zhang et al. 2021). In the federated learning approach, the confidentiality of sensitive information is maintained by not transferring any data to a central server for training and instead sharing only locally trained models. However, concerns remain about the potential to access critical information through the parameters of the shared models. To address these privacy concerns, various algorithms have been developed (Kaur et al. 2024). Some methods have been developed in the literature for federated learning architecture.

- Federated Averaging (FedAVG): It is a prominent method in the field of federated learning and was proposed by Brendan McMahan and colleagues in 2017. FedAVG is based on averaging the weights of models trained locally on clients on a central server. This approach eliminates the need for centralized data collection and significantly reduces communication costs while ensuring data privacy (McMahan et al. 2017).
- Federated Gradient Information Aggregation (FedGiA): It aims to solve the communication overhead and statistical heterogeneity problems with a hybrid approach (Zhou and Li 2023).
- Federated Asynchronous Update (FedAvu): It addresses the problem of system heterogeneity by enabling participants to work at their own speed (Sun et al. 2023).
- Fair Federated Learning (FairFed): Addresses the problem of statistical heterogeneity by focusing on ensuring balanced model performance across different groups (Ezzeldin et al. 2023).
- Privacy-oriented Robustness of Secure Federated Learning (RoFL): It attempts to ensure data privacy and security through secure gradient aggregation and encryption techniques (Lycklama et al. 2023).

For the federated learning architecture, duCBA method will be used in this study. duCBA is a method for aggregating models trained with the Classification Based

on Association Rules (CBA) algorithm (Büyüktanir et al. 2023). In this process, the support and confidence values of each rule are recalculated and the rules are ranked. For rules with the same labelling from different models, the presence rate of each rule in the dataset is checked and the support and confidence values are updated accordingly. For rules with the same content but different labels, the support value is checked and the rule with the higher support value is included in the list. The rule list formed at the end of this process forms the basis of the final model. duCBA algorithm is developed based on the relational classification approach and eliminates the need for data to be trained on local devices and collected on a central server. In this way, network traffic and data transfer are minimized while data confidentiality is maintained.

Federated learning is an innovative approach that raises privacy and security standards in the processing of medical data. This method processes data on local devices, eliminating the need to move patient information to centralized servers. Thus, while protecting individual data privacy, the risk of patient information being affected by possible security vulnerabilities is minimized. Many studies demonstrating the effectiveness of federated learning in healthcare emphasize the potential benefits of this approach. Some of these studies are summarized below:

- Differential privacy-enabled federated learning for sensitive health data (2019): Choudhury et al. stated that federated learning models supported by differential privacy enable local processing of sensitive health data, thus eliminating the need for data collection in centralized systems and significantly reducing the risk of leakage of patient data (Choudhury et al. 2019).
- Exploratory Analysis of Federated Learning Methods with Differential Privacy on MIMIC-III (2023): Horvath et al. examined the effectiveness of differential privacy and federated learning on the MIMIC-III dataset. In this research, it is stated that successful models can be developed while maintaining the privacy of patient data and medical data can be shared securely with distributed learning processes (Horvath et al. 2023).
- Federated learning model for healthchain system (2021): Durga and Poovammal discussed the Healthchain system in which the federated learning model is integrated with blockchain-based security measures and emphasized that this system provides secure data processing and sharing without the need for a centralized intermediary (Durga and Poovammal 2021).
- Comparison of machine learning algorithms and feature visualization analysis for diabetes risk prediction (2023): Compared the accuracy performance of machine learning algorithms in diabetes prediction and found that some algorithms are more successful in risk prediction and feature visualization analysis supports this success (Chen 2023).
- Diabetes type 2 classification using machine learning algorithms with up-sampling technique (2023): Demonstrated that using up-sampling technique to balance unbalanced datasets in diabetes classification improves the accuracy of machine learning algorithms (Hama Saeed 2023).
- Heart Disease Prediction Using Machine Learning (2022): In heart disease prediction, he achieved high accuracy rates with algorithms such as Random Forest, Logistic Regression and Support Vector Machines, and stated that these models are effective in analysing risk factors such as age, blood pressure and cholesterol levels (Kumar and Choudhury 2022).
- Heart Disease Prediction Using Federated Learning (2024): Predicted heart disease in a federated learning environment using datasets from the UCI Machine Learning Repository and reported high accuracy in predicting cardiovascular risk factors (Gupta et al. 2024).
- Stroke Prediction with Federated Learning (2025): This study proposes an ANN-based federated learning architecture using real stroke case data within the scope of Healthcare 4.0. The privacy-preserving FL framework, integrated with a 5G communication channel, demonstrates suitability for real-time deployment on healthcare wearable devices. The results show that the proposed model achieves 5–10% higher accuracy compared to traditional approaches, along with strong precision and recall performance (Bhatt et al. 2025).

These studies show that federated learning has an important application area in healthcare and an active research process continues in this field. In particular, by seeking solutions to critical problems such as data privacy, security and system efficiency, federated learning enables the training of powerful and effective machine learning models on medical data. In this context, the existing literature demonstrates the potential of federated learning to preserve privacy and efficiency on medical data. In the literature, federated learning methods have generally addressed critical problems such as data privacy and system efficiency when performing decentralised data analysis. However, while existing methods mostly address the problems of heterogeneity in model updates,

communication overhead and statistical imbalance, they have not focused on a federated learning architecture based on associative classification. duCBA aggregation algorithm offers a model focused on associative classification, especially by integrating the CBA algorithm into the federated learning environment. This algorithm aims to create a more optimized and unified model by updating the support and confidence values of rules from different clients. The generated rules have a structure that can be both interpreted and analyzed. In this way, it is possible not only to evaluate current situations, but also to predict future trends and risks. The gap in the literature is in updating relational classification rules in accordance with the federated learning architecture and sharing them while maintaining confidentiality. duCBA fills this gap and offers an innovative solution that both prioritizes confidentiality and minimizes data transfer. In this respect, unlike other methods, it aims to increase both model accuracy and system efficiency.

3. Materials and Methods

In this section of the research, the dataset used for model training, the technologies and methods used for the developed application are explained in detail.

3.1 Datasets

In the conduct of this research, the 'Diabetes Health Indicators Dataset' (Diabetes Health Indicators Dataset 2024) and 'Heart Attack Dataset' (Fayez 2024) datasets available on the Kaggle platform were used to train and evaluate the models. In the model training process, 20% of each dataset was reserved for testing and the remaining 80% was used for training. Three different feature selection algorithms, Principal Component Analysis (PCA), Chi-Square and Mutual Information algorithms, which are common and effective in the literature, were applied on these datasets. PCA focuses on dimension reduction and analyzing the variance of features. Chi-Square selects significant features on the basis of statistical independence test. Mutual Information evaluates the strength of the relationship between the target variable and the features by providing an information theory-based approach. The algorithms used were chosen because they represent different perspectives of feature selection techniques. This diversity provided a balanced assessment of the impact of different methods on the model.

3.1.1 Diabetes Health Indicators Dataset

The Diabetes Health Indicators Dataset is based on data from the Behavioural Risk Factor Surveillance System (BRFSS) survey conducted by the US Centers for Disease Control and Prevention (CDC). This dataset contains information on diabetes and other factors that may affect the general health of individuals. The dataset is balanced within both classes and consists of 70.692 rows and 22 columns. The dependent variable 'Diabetes_binary' column provides binary information about the respondent's diabetes status. The other 21 independent variables cover a variety of factors representing individuals' lifestyles, health behaviors and general health status. These variables include high blood pressure, cholesterol level, body mass index (BMI), physical activity level, smoking, history of heart disease, general health status and income levels. After analysing the data series, it was decided not to perform any preprocessing on the dataset.

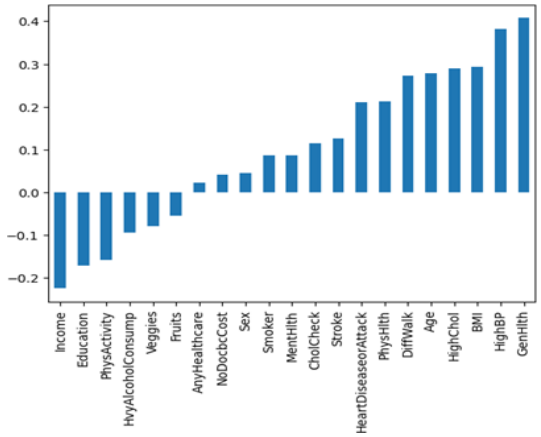

Figure 2. Correlation of Independent Variables with the Diabetes_binary

Figure 2 shows the independent variables in the dataset and their correlation graph on diabetes. According to this graph, the effect of independent variables on diabetes can be positive or negative. Positive values increase the risk of diabetes, while negative values decrease the risk. According to the Figure 2, 'GenHlth' and 'HighBP' are strongly associated with diabetes, which may increase the risk of diabetes. On the other hand, variables such as 'Income' are negatively associated with diabetes risk, suggesting that a higher income level may reduce the risk of diabetes.

As shown in Figure 3, Mutual Information scores, which represent the information sharing of features with the target variable, were evaluated. GenHlth, HighBP, and BMI show a strong information sharing with the target variable with high scores. On the other hand, CholCheck, Smoker, Stroke, PhysActivity, Fruits, Veggies, HvyAlcoholConsump, AnyHealthcare, NoDocbcCost, MentHlth, Sex, Education have low scores. This indicates that these characteristics do not have a significant effect on the target variable and can be removed from the model.

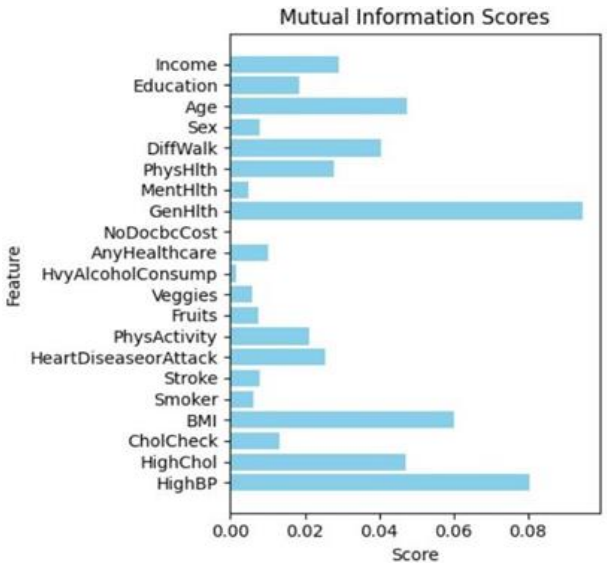

Figure 3. Mutual Information Scores for Diabetes_binary

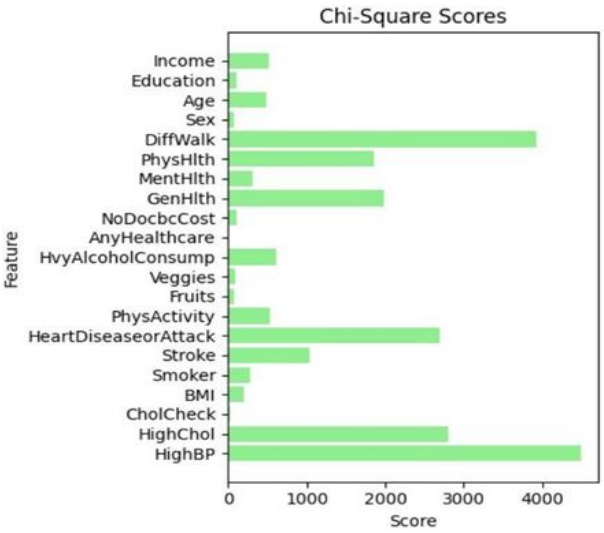

Figure 4. Chi-Square Scores for Diabetes_binary

As shown in Figure 4, Chi-Square analysis reveals the dependency relationship of the categorical variables with the target variable. HighBP, DiffWalk, and HeartDiseaseorAttack have the highest Chi-Square scores and show a strong relationship with the target variable. However, features with low scores such as CholCheck, Veggies, Fruits, AnyHealthcare and Sex show a weak relationship with the target variable. This indicates that these characteristics do not have a significant effect on the target variable and can be removed from the model.

As shown in Figure 5, PCA analysis evaluates the contribution of each attribute in explaining the total variance. Fruits, MentHlth, and Veggies make the highest contribution in PCA. Since the contribution of all features is significant, there is no feature that should be removed according to the PCA results. PCA analysis provided important information in terms of dimension reduction and variance explanation, but there was no need to remove any trait in this dataset.

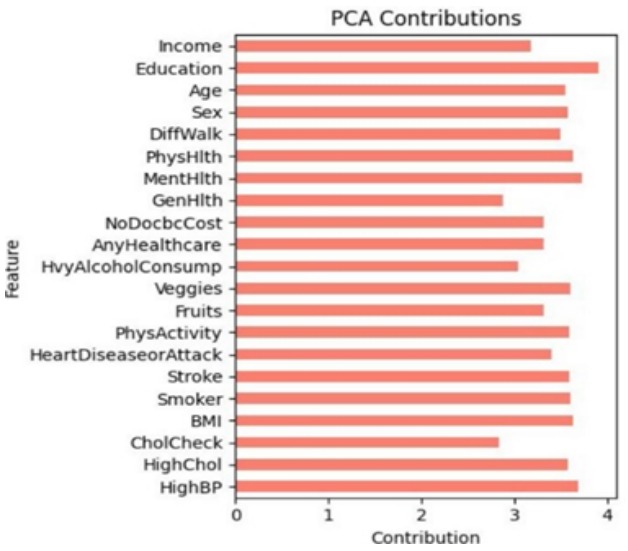

Figure 5. PCA Contributiond for Diabetes_binary

Table 1. Feature Selection of Diabetes Health Indicators Dataset

Features	Mutual Information	Chi Square	PCA
HighBP	+	+	+
HighChol	+	+	+
CholCheck	-	-	+
BMI	+	+	+
Smoker	-	+	+
Stroke	-	+	+
HeartDiseaseorAttack	+	+	+
PhysActivity	-	+	+
Fruits	-	-	+
Veggies	-	-	+
HvyAlcoholConsump	-	+	+
AnyHealthcare	-	-	+
NoDocbcCost	-	+	+
GenHlth	+	+	+
MentHlth	-	+	+
PhysHlth	+	+	+
DiffWalk	+	+	+
Sex	-	-	+
Age	+	+	+
Education	-	+	+
Income	+	+	+

Table 1 shows which features were selected as a result of the three different feature selection methods used on the Diabetes Health Indicators Dataset.

3.1.1 Heart Attack Dataset

In this research, the Heart Attack Dataset obtained from Kaggle was used. The dataset consists of 319.795 rows and includes HeartDisease, BMI, Smoking, Alcohol Drinking, Stroke, PhysicalHealth, MentalHealth, Walking Differently, Gender, AgeCategory, Diabetic, PhysicalActivity, GeneHealth, SleepTime, Asthma, KidneyDisease and SkinCancer. The target variable indicates whether the individual has heart disease (Yes: Healthy, No: Disease present). These features provide a meaningful and rich data source for the prediction of

heart disease. Before using this dataset, we performed the following preprocessing operations on the dataset.

- Changed all yes values in the dataset to 1 and no values to 0.
- Changed all male values in the dataset to 1 and female values to 0.
- The AgeCategory property was re-categorised in the range 0-12 (0 = 18-24, 12 = 80 or older).
- The GenHealth property was re-categorised in the range 0-4 (0 = poor, 4 = excellent).
- Race feature removed.
- The original Heart Attack dataset exhibits a substantial class imbalance. To prevent the model from leaning toward the majority class, random under sampling was applied. After this balancing step, the 0 and 1 classes of the HeartDisease variable were equalized, resulting in a balanced dataset of 52.952 records, with 26.476 samples in each class.

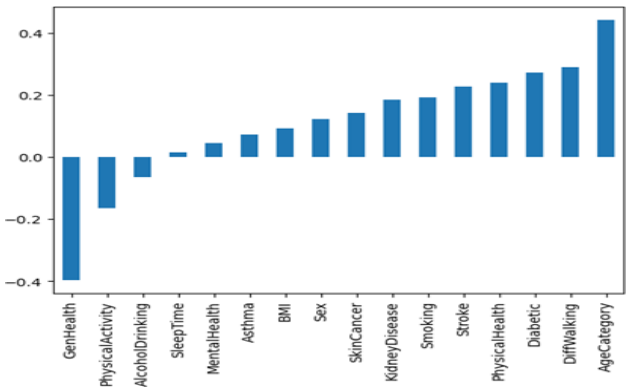

Figure 6. Correlation of Independent Variables with the HeartDisease

The correlation graph in Figure 6 shows the correlation between the variable 'HeartDisease' and other variables. The highest positive correlation is seen with the age category, indicating that the risk of heart disease increases with increasing age. Factors such as diabetes, smoking, history of stroke and BMI are also among the important variables that increase the risk of heart disease by showing positive correlations. On the other hand, general health status and physical activity are negatively correlated, meaning that as these variables increase, the risk of heart disease decreases. In addition, variables such as sleep duration and mental health appear to have weaker associations.

As shown in Figure 7, Mutual Information scores, which measure the information sharing of the features with the target variable, are analyzed. This analysis reveals that variables such as age category and diabetes are the best predictors of the target variable. Variables such as general health status, physical health, difficulty walking and history of stroke also make notable contributions, while factors such as smoking and sleep duration appear to be less influential. This type of analysis can help to identify which variables should be prioritized in modelling and predicting processes. As a result of the analysis, BMI, AlcoholDrinking, MentalHealth and Asthma variables with low information scores on the target variable were excluded from the model.

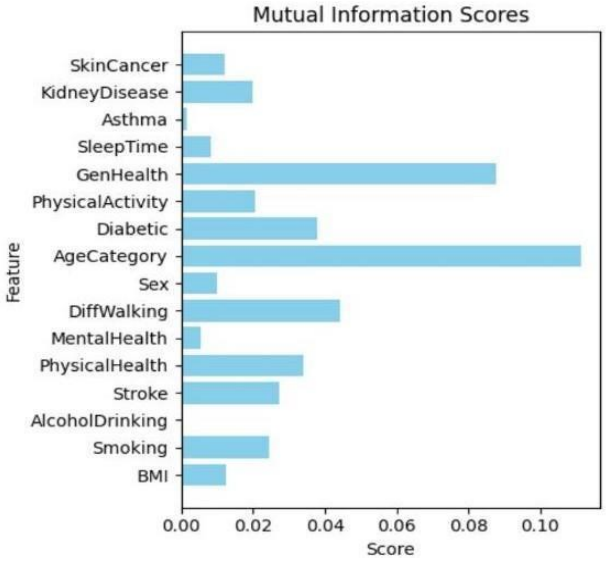

Figure 7. Mutual Informatıon Scores for HeartDisease

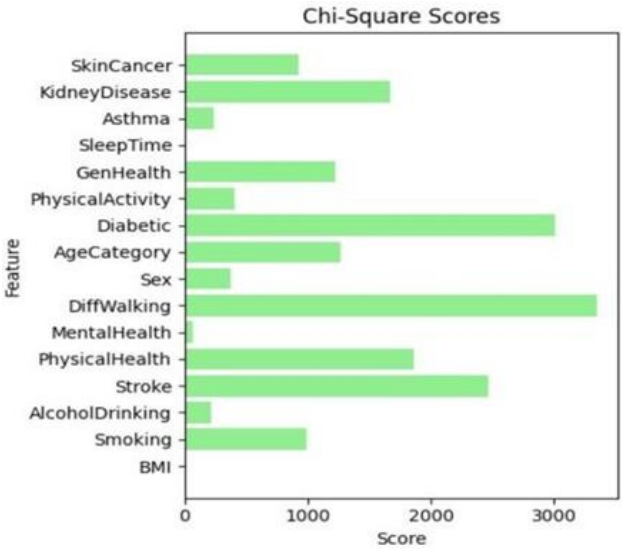

Figure 8. Correlation of Independent Variables with the HeartDisease

As shown in Figure 8, Chi-Square Scores, which measure the information sharing of the features with the target variable, are analyzed. As a result of the analysis, BMI, MentalHealth and SleepTime variables, which have low Chi-Square scores on the target variable, were excluded from the model. Since these variables did not make a significant contribution to the estimation of the target variable, it was deemed appropriate to remove them from the dataset in order to improve the performance of

the model and reduce unnecessary complexity. This process allowed the model to focus on variables with stronger relationships, allowing for a more effective and meaningful analysis. As a result of the analysis, it was determined that especially AgeCategory, DiffWalking, and Diabetic variables had strong effects on the target variable and the model focused on these variables.

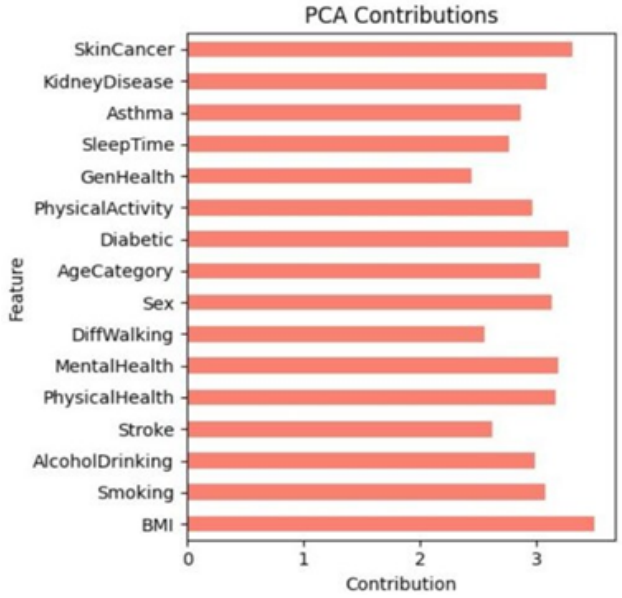

Figure 9. PCA Contributiond for HeartDisease

Table 2. Feature Selection of Heart Attack Dataset

Features	Mutual Information	Chi Square	PCA
BMI	-	-	+
Smoking	+	+	+
AlcoholDrinking	-	+	+
Stroke	+	+	+
PhysicalHealth	+	+	+
MentalHealth	-	-	+
DiffWalking	+	+	+
Sex	+	+	+
AgeCategory	+	+	+
Diabetic	+	+	+
PhysicalActivity	+	+	+
GenHealth	+	+	+
SleepTime	+	-	+
Asthma	-	+	+
KidneyDisease	+	+	+
SkinCancer	+	+	+

As shown in Figure 9, PCA which measure the information sharing of the features with the target variable, are analyzed. As a result of the analysis, no variable was omitted. The results show that especially AgeCategory, Diabetic, and DiffWalking variables are critical for the overall success of the model. Although other variables contributed less, it is important to include all of them in order to preserve the integrity of the data and to provide a broader analysis perspective. These results provide useful clues on how to structure the model based on PCA's variance explanations considering all variables.

Table 2 shows which features were selected as a result of the three different feature selection methods used on the Heart Attack Dataset.

3.2 Prototype Application

Three different environments were used in the prototype application developed within the scope of the research which are Clients, API and Central Server (Demir et al. 2024). Figure 10 shows the working principle of the prototype application. In the first step, local clients train models on their own data using the CBA method and send these models to the API server. The API server collects the local models and forwards them to the central server. The central server creates an updated model by aggregating the incoming models with the duCBA federated aggregation algorithm. This updated model is sent back to the API server. The API server distributes this updated model to the clients, enabling each client to use a more robust and global model. Figure 10 shows this process as 1 round. Due to the structure of federated learning, the tour continues continuously and models trained with more data are trained. This process provides an approach that aims to learn on local data and train a global model in this way, while maintaining the confidentiality of the data.

In this study, there were no physically distributed clients or institutions available from real-world environments. Therefore, the federated learning process was simulated using a setup designed to closely mimic real distributed scenarios. Because of the structure of this simulation, each client trained its model locally and sent it to the server only once, allowing a single federated round to be completed through model aggregation. As a result, unlike standard FL implementations, the system did not require multiple iterative training rounds (Ma et al. 2023).

In practical federated learning applications, stopping criteria (Zhang et al. 2021) can be defined in various ways such as using a fixed number of rounds, reaching a target accuracy, detecting convergence through minimal parameter changes, reducing global loss below a certain threshold, or meeting time and resource limits. However, the goal of this research was to directly compare aggregation methods like duCBA and FedAvg. For this reason, multi-round stopping strategies were not included in the experimental setup. Accordingly, the stopping criterion in this study was simply the successful aggregation of local client models, followed by evaluating the performance of the resulting global model.

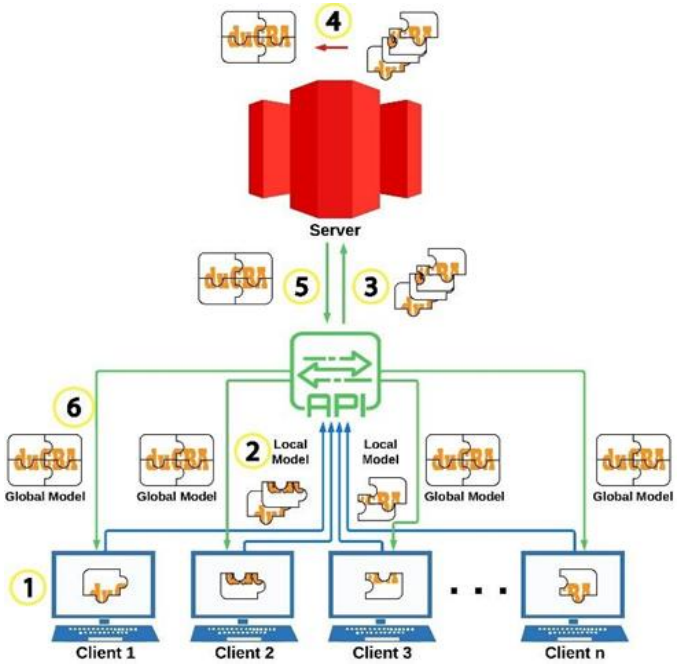

Figure 10. duCBA Workflow Diagram

The study focused mainly on comparing the performance of different federated aggregation algorithms. For this reason, all experimental procedures were applied consistently across each method. As part of the comparison involving the FedAvg algorithm, Artificial Neural Network (ANN) models were trained on the client side, and the server aggregated their weights using the standard FedAvg mechanism.

3.2.1 Clients

Identical virtual machines were created on Google Cloud Platform for clients. The e2-medium machine type using x86/64 architecture was used. Python version 3.8.10 and numpy, pandas, scikit-learn, pickle, pyfim, requests, pyarc libraries were used for model training. In the model training process, 20% of the datasets used in the model training process were reserved for testing and the remaining part was distributed equally to the number of clients. CBA algorithm was used for model training. The support and confidence values we will use for CBA are set as 0.2 and 0.5 respectively. These values are pre-tested and approved values. The models created and trained with these settings are sent to the API in json format via HTTP request. API sends these models to the central server for aggregating. The models aggregated on the central server are sent to the clients by the API and the clients continue their training with the global model (Demir et al. 2024).

3.2.2 API

For the API, e2-medium machine type was used on Google Cloud Platform. Python version 3.8.10 and flask, numpy, pandas, scikit-learn, pickle, pyfim, websockets, requests, pyarc libraries were used. Requests to the API were received using port 5000. In the prototype application tests, the models coming from clients were collected 10 times with the API and sent to the central server. These models, which were sent to the central server by websockets, were received from the central server and distributed to the clients by API (Demir et al. 2024).

3.2.3 Central Server

For the central server, e2-medium machine type was used on Google Cloud Platform. Python version 3.8.10 and numpy, pandas, scikit-learn, pickle, pyfim, websockets, requests, pyarc libraries were used. Websocket connection was realized using port 8000. The duCBA federated aggregation algorithm was used to aggregate the models coming from the API. The algorithm aims to aggregate local models trained with the CBA method on each client on the server. During the aggregating process, the rules merged from each client are thoroughly evaluated and aggregated by recalculating the support and confidence values (Büyüktanır et al. 2023). This evaluation process is performed in order to achieve an optimal result when similar or identical rules come from different clients. During aggregation, the support and confidence values are updated by checking how many rules marked with the same label but coming from different clients are present in the dataset. Thus, the validity of each rule for the whole system is calculated more accurately. The frequency of occurrence (support value) is compared between rules with the same content

but marked with different labels, and the rule with the higher support value is selected and added to the final list. After the update, the rules are sorted according to the new confidence values and the support value is taken into account in case of equal confidence values. If the support values are the same, the first rule in the list is ranked higher (Demir et al. 2024).

This comprehensive evaluation and ranking process results in the creation of a aggregated final model, where all rules contributed by the models from the clients are integrated. The local models created by the clients via the API on the central server are aggregated using the duCBA federated aggregation algorithm, as previously mentioned. The resulting global model is then sent back to the clients through the API, completing the federated learning cycle (Demir et al. 2024).

4. Experimental Results

First of all, we compared two different ways of training a model in distributed settings. First, we looked at the non-federated approach, where all client data is collected on a central server and a single CBA model is trained there. Then we examined the federated approach, where each client trains its own CBA model locally. These local models are then sent to the server and combined using our duCBA method. To make a fair comparison, we ran experiments on each dataset using a setup with three clients.

After comparing the federated and non-federated training setups, we also evaluated the performance of different aggregation methods. We compared duCBA with FedAVG, one of the most widely used aggregation algorithms in federated learning. duCBA is a relational classification-based aggregation algorithm designed for federated learning architectures (Büyüktanır et al. 2023) For this reason each client trains a CBA model locally and sends it to the server, where duCBA merges all client-side models into a global model. For the FedAVG based setup, clients trained their local models using an ANN, and the server generated the global model by applying the FedAVG aggregation algorithm. The ANN architecture was selected for FedAVG because it offers a lightweight and computationally efficient model suitable for federated environments. Its simple structure also makes it possible to evaluate the performance of FedAVG more clearly, without the influence of unnecessary model complexity (Bhatt et al. 2024).

We evaluated both aggregation strategies on two datasets while running the experiments with 2, 4, and 8 clients. The comparison focused on standard performance metrics; precision, recall, F1-score, and accuracy. During the tests, we also applied three different feature selection methods to observe how they influenced the results.

The architecture and hyperparameters of the ANN model were determined through an iterative trial and error process, drawing on widely accepted practices in the machine learning literature for binary classification tasks. In the first layer of the ANN, the input features are transformed non-linearly using a Dense layer with 64 neurons and a ReLU activation function, defined according to the input dimension. The output layer consists of a single neuron with a sigmoid activation function, producing the model's prediction as a probability between the two classes. For optimization, the model employed the Adam optimizer with a learning rate of 0.001. Classification errors were minimized using the binary_crossentropy loss function, and performance was evaluated with the accuracy metric. The model was trained for a total of 5 epochs using mini-batches of 32 samples in each iteration. All models were trained for 10 independent iterations to analyze the effect of the number of clients, feature selection methods, and aggregation algorithms on overall performance.

Table 3. Comparative Classification Results for Centralized and Federated Models

	Diabetes Health Indicators Dataset		**Heart Attack Dataset**	
	CBA	**duCBA**	**CBA**	**duCBA**
Precision	0.7	0.74	0.69	0.77
Recall	0.7	0.74	0.69	0.77
F1-Score	0.70	0.74	0.69	0.77
Accuracy	0.70	0.74	0.69	0.77
Time	234.20	228.75	92.45	90.35

Table 3 shows how the models performed when trained in both federated and non-federated settings. We use these results as a benchmark to more clearly compare the two main federated aggregation methods evaluated in our study: duCBA and FedAVG.

The model merges with duCBA for the Diabetes dataset showed improved performance compared to the model trained with centralized CBA. All evaluation metrics increased from 0.70 to 0.74. This difference is even more pronounced in the Heart Attack dataset. The federated duCBA model reached 0.77 across all metrics, while the centralized CBA model remained at 0.69. When the average execution times were examined (in seconds), the centralized CBA model required 234.20 seconds for the first dataset and 92.45 seconds for the second. In contrast, the federated duCBA model completed the same tasks slightly faster, finishing in 228.75 seconds and 90.35 seconds, respectively.

This demonstrates that the federated learning scenario produces more successful models than the non-federated scenario, while also providing data privacy and delivering a more reliable model. These results demonstrate that the federated learning architecture not only protects data privacy but also produces more successful models compared to a centralized (non-federated) structure. Furthermore, it is evident that duCBA operates more efficiently in terms of computation times and offers a scalable solution.

Table 4. Diabetes Health Indıcators Dataset Classification Report

		Mutual Information				Chi Square				PCA			
		Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
2	duCBA	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75	0.75
	FedAVG	0.75	0.68	0.66	0.68	0.75	0.71	0.70	0.71	0.75	0.74	0.74	0.74
4	duCBA	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.69	0.69	0.69
	FedAVG	0.75	0.74	0.74	0.74	0.75	0.74	0.74	0.74	0.75	0.74	0.74	0.74
8	duCBA	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
	FedAVG	0.74	0.73	0.73	0.73	0.74	0.74	0.74	0.74	0.72	0.72	0.72	0.72

Table 5. Heart Attack Dataset Classification Report

		Mutual Information				Chi Square				PCA			
		Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score	Accuracy
2	duCBA	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78	0.78
	FedAVG	0.78	0.78	0.78	0.78	0.79	0.79	0.79	0.79	0.79	0.79	0.79	0.79
4	duCBA	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70	0.70
	FedAVG	0.75	0.75	0.75	0.75	0.74	0.74	0.74	0.74	0.75	0.75	0.75	0.75
8	duCBA	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69	0.69
	FedAVG	0.75	0.75	0.75	0.75	0.76	0.75	0.75	0.75	0.75	0.74	0.74	0.74

Table 4 shows the results obtained using the Diabetes Health Indicators Dataset as the dataset for model training. These results compare the performance of three different feature selection algorithms and two different aggregation algorithms with 2, 4 and 8 clients. In general, FedAVG stands out with higher F1-Score and Precision values, especially as the number of clients increases, while duCBA provides more stable and balanced results with fewer clients. These results show that duCBA can provide stable learning in small and medium-sized systems. In particular, similar results are obtained with different feature selection algorithms such as Mutual Information, Chi Square and PCA, indicating that the method is compatible with various data preprocessing methods. However, although the increase in the number of clients causes small fluctuations in some cases, duCBA stands out as a reliable method in distributed systems requiring data privacy by providing constant values between 0.69-0.75 in terms of overall accuracy. These results show that duCBA is a suitable method for applications that require more stable and reliable performance.

Table 5 shows the results obtained using the Heart Attack Dataset as the dataset for model training. These results compare the performance of three different feature selection algorithms and two different aggregation algorithms with 2, 4 and 8 clients. In general, FedAVG stands out with higher F1-Score and Precision values, especially with the increase in the number of clients, while duCBA provides balanced and consistent results with fewer clients. duCBA showed similar performance with different feature selection algorithms such as Mutual Information, Chi Square and PCA, and it was observed that the effect of data preprocessing algorithms used with this method was limited. Although the increase in the number of clients caused fluctuations in some metrics, the results of duCBA in terms of accuracy remained stable between 0.78-0.69, indicating that the method provides a reliable solution for small and medium-sized systems.

Table 6 shows, when comparing the average execution times, the duCBA algorithm is observed to offer significantly shorter durations than the FedAVG algorithm on both datasets. For example, the duration for duCBA on the Diabetes Health Indicators Dataset with 2 clients is 84.37, while the duration for FedAVG is 430.68. As the number of clients increases, a slight downward trend in duCBA's execution time has been observed, indicating that duCBA is efficient in terms of scalability in distributed systems. Although the execution time also decreased for FedAVG as the number of clients increased, its absolute

values remained significantly higher than duCBA's durations.

Table 6. Average Execution Time (seconds)

Number of Clients	Aggregation Algorithm	Diabetes Health Indicators Dataset	Heart Attack Dataset
2	duCBA	84.37	90.35
	FedAVG	430.68	306.45
4	duCBA	80.76	82.13
	FedAVG	299.02	233.22
8	duCBA	78.78	77.12
	FedAVG	267.50	220.62

In conclusion, these findings demonstrate that duCBA offers a competitive or even superior speed compared to centralized systems, even when operating in a privacy-preserving federated environment. This positions duCBA as a strong alternative for applications targeting reliability and stability, particularly in scenarios involving a low number of clients and sensitive data.

When the results obtained are evaluated, it is seen that the duCBA algorithm provides stable and balanced results especially in small and medium-sized systems. duCBA has proved to be an effective method in scenarios requiring data privacy by exhibiting a stable and reliable performance without fluctuations depending on the number of clients. Furthermore, the consistent results regardless of the feature selection algorithms show that the method works well with different data preprocessing methods. While FedAVG provided superior performance in the F1-Score and Precision metrics in systems with more clients, duCBA is more effective in low-client and sensitive data scenarios. These findings suggest that duCBA is a strong alternative for applications that aim for reliability and stability in distributed systems.

5. Discussion

In this study, model training was performed using a federated learning architecture with two different datasets. Two different federated aggregation algorithms were used to aggregate the models trained on the clients to the server. The first, FedAVG, is a method commonly used in federated systems that aggregates the model parameters obtained from each client by averaging them. The other is the duCBA algorithm. This aggregation algorithm is a federated aggregation method based on relational classification. It aggregates the models trained with CBA on each client at the server. The data sent to the server are actually relational rules, and these rules are aggregated based on support and confidence parameters. In the duCBA method, each step in the model aggregation process is clearly defined and the principles of how the model works are presented in an understandable way. This shows that the operation of the model and the decisions it makes can be transparently explained, contributing to the field of Explainable Artificial Intelligence (XAI) (Kalasampath et al. 2025). However, the final model obtained with the FedAVG federated aggregation algorithm alone is not sufficient in terms of explainability.

The core argument of this study is that duCBA not only offers a model that is competitive in terms of performance but also provides a level of transparency that traditional weight-aggregation methods like FedAVG inherently cannot. While FedAVG optimizes the weights of the final model, it is nearly impossible to find a clinical or semantic equivalent for these weights or to explain why the model made a specific prediction. In contrast, duCBA's aggregation of relational rules allows for the direct examination and interpretation of the model's decision logic. This distinction is critically important, particularly in sensitive domains such as healthcare, where interpretability is paramount.

In this context, in the final model obtained with the duCBA federated aggregation algorithm, the strongest rules were selected based on the support and confidence values calculated separately for each dataset, and these rules were interpreted in detail. The interpretations of the rules obtained for the Diabetes Health Indicators Dataset are as follows:

1. Health Conditions and Diabetes: In the rules indicating situations where the risk of diabetes is low, it is generally observed that healthy individuals are free from major health issues:

- High Blood Pressure (HighBP = 0.0): Maintaining normal blood pressure levels stands out as an important factor in reducing the risk of diabetes.
- High Cholesterol (HighChol = 0.0): Low cholesterol levels indicate that metabolism is functioning healthily, indirectly suggesting a reduced risk of diabetes.
- Difficulty Walking (DiffWalk = 0.0): Preserved mobility implies that the individual may have an active lifestyle, which in turn reduces the risk of diabetes.
- Stroke (Stroke = 0.0) and Heart Problems (HeartDiseaseorAttack = 0.0): The absence of cardiovascular issues suggests a lower likelihood of diabetes-related complications.

These findings highlight that a healthy lifestyle (balanced nutrition, regular exercise, avoiding stress, etc.) can significantly reduce the risk of diabetes.

2. Physical Activity and Diet Effect: It is observed that individuals with high physical activity and vegetable consumption have a lower risk of diabetes. Sample rule: *{PhysActivity = 1.0, Veggies = 1.0, DiffWalk = 0.0, HighBP = 0.0} → Diabetes_binary = 0.0 with a confidence level of 79.4%*. This finding suggests that maintaining a healthy diet and engaging in regular physical activity are effective methods for controlling blood sugar levels. Healthy eating habits can increase insulin sensitivity and support weight management.

3. Access to Healthcare and Diabetes: It is observed that diabetes rates are lower in situations where access to healthcare and the cost of doctor visits are low. Sample rule: *{HighBP = 0.0, AnyHealthcare = 1.0, NoDocbcCost = 0.0} → Diabetes_binary = 0.*0 with a confidence level of 81.6%.
This rule emphasizes the critical role of regular health check-ups in the early diagnosis and prevention of diabetes. Additionally, it suggests that easier access to healthcare can improve individuals' overall health status.

4. The Effect of Multiple Health Factors: For example, the combination *{HighChol = 0.0, Stroke = 0.0, DiffWalk = 0.0, HeartDiseaseorAttack = 0.0, HighBP = 0.0}* has a confidence level of 84.3% for no diabetes.

This rule clearly demonstrates that when a healthy individual is free from multiple risk factors, the risk of diabetes is significantly lower. In other words, rather than focusing on a single health factor, it suggests that optimizing an individual's overall health profile is essential.

In conclusion, based on the rules obtained, the most important factors affecting diabetes are lifestyle, access to healthcare, physical activity levels, and health history. The data supports the need to increase diabetes prevention health practices and awareness. Individuals who adopt a healthy lifestyle have a higher chance of being protected from diabetes. Additionally, making healthcare accessible to everyone is a critical step in improving public health. Based on these rules, more targeted preventive policies and individual awareness programs can be developed.

The interpretations of the rules obtained for the Heart Attack Dataset are as follows:

1. Health Conditions and Heart Disease: Individuals without heart disease be associated with the absence of specific risk factors:

- Diabetes (Diabetic = 0.0): The majority of individuals without heart disease do not have diabetes. Sample rule: *{Diabetic = 0.0, DiffWalking = 0.0, KidneyDisease = 0.0, Stroke = 0.0, SkinCancer = 0.0, Smoking = 0.0} → HeartDisease = 0.0* combination is associated with the absence of heart disease at a 76.2% accuracy rate.
- Difficulty Walking (DiffWalking = 0.0): Individuals who maintain mobility have a lower risk of heart disease. This indicates the positive impact of physical activity on overall health.
- Non-Smoking (Smoking = 0.0): Individuals who do not smoke have a reduced risk of heart disease. Sample rule: *{Smoking = 0.0, DiffWalking = 0.0, Diabetic = 0.0, Stroke = 0.0} → HeartDisease = 0.0* combination has an accuracy rate of 76.8%.

2. Physical Activity and Health: The positive effect of physical activity on heart health is clearly visible in the dataset. Sample rule: {PhysicalActivity = 1.0, KidneyDisease = 0.0, Smoking = 0.0, Diabetic = 0.0, Stroke = 0.0} → HeartDisease = 0.0 combination is associated with no heart disease at a 76.6% confidence level. This highlights the role of regular physical activity in the prevention of heart diseases.

3. The Effect of Multiple Health Factors: Individuals who are free from multiple risk factors have a significantly lower risk of heart disease. Sample rule: *{Diabetic = 0.0, DiffWalking = 0.0, KidneyDisease = 0.0, Stroke = 0.0, PhysicalActivity = 1.0, Smoking = 0.0} → HeartDisease = 0.0* combination has an accuracy rate of 74.9%. This suggests that improving the overall health profile is essential.

4.Kidney Diseases and Heart Health: Individuals without kidney disease have a lower risk of heart disease. Sample rule: *{KidneyDisease = 0.0, SkinCancer = 0.0, Smoking = 0.0, Diabetic = 0.0, Stroke = 0.0} → HeartDisease = 0.0* combination has an accuracy rate of 77.2%.

The analyses derived from the obtained rules highlight the importance of promoting physical activity, reducing smoking, controlling diabetes and kidney diseases, and maintaining mobility in preventing heart disease. Developing awareness programs and health policies targeting these factors can significantly improve the overall health level of society.

The performance of the duCBA algorithm proposed in this study was evaluated against the FedAVG architecture through the comprehensive experiments summarized in Table IV and Table V. When examining the results for the Diabetes and Heart Attack datasets, it is observed that duCBA achieves accuracy scores between 0.69 and 0.78 as the number of clients varies. Under the same conditions, the FedAVG algorithm produced relatively

more stable accuracy values, ranging from 0.72 to 0.79. This modest difference can be explained by the fact that duCBA constructs the global model through logical rule aggregation rather than by averaging numerical model weights. While deep learning-based approaches such as FedAVG can yield slightly higher accuracy by capturing more complex data representations, this comes at the cost of creating an entirely opaque, black-box model.

However, in sensitive decision support areas such as healthcare, accuracy alone is not enough; the interpretability of the decision process (XAI) is equally essential. In this context, the slight sacrifice in accuracy is justified, as duCBA provides a fully transparent and traceable decision-making framework.

Another important finding of the study relates to scalability and communication efficiency. The "Average Execution Times" reported in Table VI clearly indicate that duCBA outperforms FedAVG in both computational and communication efficiency:

- For the Diabetes dataset: With 2 clients, FedAVG required 430.68 seconds, whereas duCBA completed the process in only 84.37 seconds (~5.1× faster). Even with 8 clients, duCBA (78.78 sec) remained substantially faster than FedAVG (267.50 sec).
- For the Heart Attack dataset: In the 2-client scenario, duCBA (90.35 sec) achieved roughly a 3.4× speed improvement over FedAVG (306.45 sec).

These results show that duCBA not only provides an interpretable model but also significantly reduces communication overhead and computation time. While FedAVG involves transferring large, high-dimensional weight matrices at each round resulting in increased bandwidth and time requirements duCBA relies on compact, rule-based model updates, enabling a more efficient and scalable design. This makes duCBA particularly suitable for real-world federated systems deployed on resource constrained edge devices.

6. Conclusion

In this study federated learning architecture, developed to ensure data privacy in distributed systems, is utilized. The primary aim of the study is to evaluate the performance of the duCBA algorithm, one of the federated aggregation algorithms that performs model aggregation on the server side within the federated learning framework, on health data. To this end, experiments were conducted on two different health datasets, Diabetes Health Indicators and Heart Attack. Three different methods for feature selection, namely Mutual Information, Chi-Square, and Principal PCA, were applied to the datasets. First, two different model-training approaches in distributed environments were compared. The first was the non-federated (centralized) setup, where all client data is gathered on a central server and a single CBA model is trained on the combined dataset. The second was the federated setup, in which each client trains its own CBA model locally, after which the server collects these models and merges them using the duCBA method. To ensure a fair comparison, both datasets were evaluated using the same distributed configuration with three clients. Following the comparison of federated and non-federated training scenarios, the study then examined the performance of different aggregation algorithms. In this phase, the proposed duCBA algorithm was evaluated against FedAVG, one of the most commonly used aggregation techniques in federated learning. This comparison aims to assess the effectiveness of duCBA on health data by highlighting the differences in accuracy rates and overall performance between the two algorithms.

The results indicate that duCBA works compatibly with different feature selection methods, and the impact of these methods on the results is minimal. It was observed that the duCBA algorithm provides consistent and balanced results, particularly in systems with a small number of clients, with accuracy rates ranging from 69% to 78%. Moreover, the results obtained with the duCBA algorithm showed nearly identical performance when compared to the FedAVG algorithm, demonstrating that both algorithms can be effectively used on health datasets.

In the studies in the literature, federated learning methods typically address privacy and system efficiency issues through decentralized data analysis. However, the interpretability, understandability, and future applicability of the obtained models have not been sufficiently addressed. The duCBA aggregation algorithm aims to fill this gap by not only preserving data privacy but also prioritizing the interpretability and analyzability of the rules derived from the model.

The duCBA federated aggregation algorithm updates and ranks the support and confidence values of the rules from the clients, making the rules in the final model easily understandable for both users and experts. This enables not only the analysis of the current situation but also the prediction of future trends and risks. As seen in this study, by explaining risk factors for chronic diseases such as diabetes or heart disease in the health domain, it facilitates the adoption of preventive measures at the

individual patient level and can also be used as a decision support mechanism in healthcare planning. Furthermore, the interpretability provided by the algorithm prevents the model from remaining a black box, allowing users to make informed decisions about the model. The transparency in tracking the decision-making process of the final model demonstrates compliance with XAI principles. Thus, duCBA contributes significantly to the federated learning literature by providing a structure that allows for visualizable, interpretable, and predictive analyses.

In conclusion, the duCBA algorithm offers the opportunity to develop a robust predictive model while ensuring the privacy of health data. This algorithm stands out as an important solution, particularly in fields such as healthcare, where sensitive data needs to be processed. It is an effective solution that simultaneously meets the requirements of data privacy and transparency. Future work will focus on adapting the duCBA algorithm to more complex and multidimensional health data and ensuring that it works effectively with different types of data.

6.1 Limitations

This study has several methodological limitations. First, the experiments were conducted in a simulated environment on the Google Cloud Platform, using identical virtual machines and data that was equally distributed among clients. This idealized setup does not fully reflect the core challenges of federated learning, such as system heterogeneity and statistical heterogeneity. Second, the study was limited to a low number of clients; the scalability and performance of duCBA in much larger-scale networks involving hundreds or thousands of clients were not tested. Finally, this research is constrained to two specific health datasets; the algorithm's effectiveness on different data types, such as medical images, signals, or time-series data, has not been investigated.

6.2 Future Work

Building on the limitations identified in this study, future research will focus on narrowing the gap between simulated experiments and real-world federated learning deployments. First, we aim to assess the robustness of the duCBA algorithm under both system and statistical heterogeneity. This involves evaluating the framework on physical edge devices with different computational capacities (e.g., Raspberry Pi units, smartphones) and testing its behavior on Non-IID data partitions, which more closely mirror real-world medical environments. Second, to further validate the scalability benefits suggested by our initial findings, future experiments will be extended to large-scale networks involving hundreds or even thousands of clients. Such evaluations will enable a more comprehensive understanding of communication demands and performance characteristics in massive federated settings.

Lastly, we plan to broaden the applicability of the duCBA framework beyond traditional tabular health data. Future work will explore how the rule-based mechanism can be adapted to support various data modalities, such as discretized time-series signals or features derived from medical imaging. This will help expand the role of explainable AI within healthcare, ensuring that interpretability is preserved across diverse prediction tasks.

Declaration of Ethical Standards
The authors declare that they comply with all ethical standards.

Credit Authorship Contribution Statement
Author-1: Methodology / Study design, Software, Validation, Formal analysis, Investigation
Author-2: Software, Validation, Formal analysis, Investigation
Author-3: Writing – original draft, Writing – review and editing, Supervision
Author-4: Investigation, Writing – review and editing, Supervision
Author-5: Investigation, Writing – review and editing, Supervision

Declaration of Competing Interest
The authors have no conflicts of interest to declare regarding the content of this article.

Data Availability Statement
All data generated or analyzed during this study are included in this published article.

Acknowledgements

This research was supported by the TÜBİTAK 2209-A "University Students Research Projects Support Program."

7. References

Albshaier, L., Almarri, S. and Albuali, A., 2025. Federated Learning for Cloud and Edge Security: A Systematic Review of Challenges and AI Opportunities. *Electronics*, 14, **5**, 1019. https://doi.org/10.3390/electronics14051019

Ali, M.K., Pearson-Stuttard, J., Selvin, E. and Gregg, E.W., 2022. Interpreting global trends in type 2 diabetes complications and mortality. *Diabetologia,* 65, **1**, 3-13. https://doi.org/10.1007/s00125-021-05585-2

Bian, J., Arafat, A.A., Xiong, H., Li, J., Li, L., Chen, H., Wang, J., Dou, D. and Guo, Z., 2022. Machine Learning in Real-Time Internet of Things (IoT) Systems: A Survey. *IEEE Internet of Things Journal*, 9, **11**, 8364-8386. https://doi.org/10.1109/JIOT.2022.3161050

Bhatt, H., Jadav, N.K., Kumari, A., Gupta, R., Tanwar, S., Polkowski, Z., Tolba, A. and Hassanein, A.S., 2024. Artificial neural network-driven federated learning for heart stroke prediction in healthcare 4.0 underlying 5G. *Concurrency and Computation: Practice and Experience*, 36, **3**, e7911. https://doi.org/10.1002/cpe.7911

Buyuktanir, B., Altinkaya, Ş., Karatas Baydogmus, G. and Yildiz, K., 2025. Federated learning in intrusion detection: advancements, applications, and future directions. *Cluster Computing*, 28, **7**, 473. https://doi.org/10.1007/s10586-025-05325-w

Büyüktanir, B., Çıplak, Z., Çil, A. E., Yakar, Ö., Adoum, M.B., & Yıldız, K. 2025. DDoS_FL: Federated learning architecture approach against DDoS attack. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 31, **6**, 1004-1018. https://dergipark.org.tr/en/pub/pajes/article/1813231

Büyüktanır, B., Yıldız, K., Ülkü, E.E. and Büyüktanır, T., 2023. du-CBA: Data-agnostic and incremental classification-based association rules extraction architecture. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 38, **3**, 1919-1929. https://doi.org/10.17341/gazimmfd.1067527

Chen, S., 2023. Comparison of machine learning algorithms and feature visualization analysis for diabetes risk prediction. *Journal of Physics: Conference Series*, 2646, **1**, 012013. https://doi.org/10.1088/1742-6596/2646/1/012013

Choudhury, O., Gkoulalas-Divanis, A., Salonidis, T., Sylla, I., Park, Y., Hsu, G. and Das, A., Differential Privacy-enabled Federated Learning for Sensitive Health Data. https://doi.org/10.48550/arXiv.1910.02578

Çıplak, Z., Yıldız, K. and Altınkaya, Ş., 2025. FEDetect: A Federated Learning-Based Malware Detection and Classification Using Deep Neural Network Algorithms. *Arabian Journal for Science and Engineering*, 50, **19**, 16107-16134. https://doi.org/10.1007/s13369-025-10043-x

Demir, A., Kulaksız, A.Y., Büyüktanır, B., Baydoğmuş, G.K. and Yıldız, K., (2024). Model Training and Real World Analysis Using Health Data with Federated Learning. 2024 International Open Source Conference (UAKK). İstanbul, Türkiye, 103-108.

Dian, F.J., Vahidnia, R. and Rahmati, A., 2020. Wearables and the Internet of Things (IoT), Applications, Opportunities, and Challenges: A Survey. *IEEE Access,* **8**, 69200-69211. https://doi.org/10.1109/ACCESS.2020.2986329

Durga, R. and Poovammal, E., (2021). Federated Learning Model for Healthchain System. 2021 6th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE). Kedah, Malezya, 1-6.

Ezzeldin, Y.H., Yan, S., He, C., Ferrara, E. and Avestimehr, A.S., 2023. FairFed: Enabling Group Fairness in Federated Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37, **6**, 7494-7502. https://doi.org/10.1609/aaai.v37i6.25911

Gupta, S., Kumar, P., Srivastava, N.V., Kumar, A. and Chaurasia, B.K., (2024). Heart Disease Prediction Using Federated Learning. Proceedings of International Conference on Recent Innovations in Computing. Singapore, 35-47.

Hama Saeed, M.A., 2023. Diabetes type 2 classification using machine learning algorithms with up-sampling technique. *Journal of Electrical Systems and Information Technology*, 10, **1**, 8. https://doi.org/10.1186/s43067-023-00074-5

Horvath, A.N., Berchier, M., Nooralahzadeh, F., Allam, A. and Krauthammer, M., 2023. Exploratory Analysis of Federated Learning Methods with Differential Privacy on MIMIC-III. *arXiv preprint arXiv:2302.04208*. https://doi.org/10.48550/arXiv.2302.04208

Kalasampath, K., Spoorthi, K.N., Sajeev, S., Kuppa, S.S., Ajay, K. and Maruthamuthu, A., 2025. A Literature Review on Applications of Explainable Artificial Intelligence (XAI). *IEEE Access*, 13, 41111-41140. https://doi.org/10.1109/ACCESS.2025.3546681

Kaur, H., Rani, V., Kumar, M., Sachdeva, M., Mittal, A. and Kumar, K., 2024. Federated learning: a comprehensive review of recent advances and applications. *Multimedia Tools and Applications*, 83, **18**, 54165-54188. https://doi.org/10.1007/s11042-023-17737-0

Kucur, E. N., Buyuktanir, T., Ugurelli, M., & Yildiz, K. (2025). Privacy-Preserving Machine Learning Techniques: Cryptographic Approaches, Challenges, and Future Directions. Applied Sciences, 16(1), 277. https://doi.org/10.3390/app16010277

Li, T., Sahu, A.K., Talwalkar, A. and Smith, V., 2020. Federated Learning: Challenges, Methods, and Future Directions. *IEEE Signal Processing Magazine*, 37, **3**, 50-60. https://doi.org/10.1109/MSP.2020.2975749

Li, M., Xu, P., Hu, J., Tang, Z. and Yang, G., 2025. From challenges and pitfalls to recommendations and opportunities: Implementing federated learning in healthcare. *Medical Image Analysis*, 101, 103497. https://doi.org/10.1016/j.media.2025.103497

Lycklama, H., Burkhalter, L., Viand, A., Küchler, N. and Hithnawi, A., (2023). RoFL: Robustness of Secure Federated Learning. 2023 IEEE Symposium on

Security and Privacy (SP). San Francisco, USA, 453-476.

Ma, Y., Woods, J., Angel, S., Polychroniadou, A. and Rabin, T., (2023). Flamingo: Multi-Round Single-Server Secure Aggregation with Applications to Private Federated Learning. 2023 IEEE Symposium on Security and Privacy (SP). San Francisco, USA, 477-496

McMahan, B., Moore, E., Ramage, D., Hampson, S. and Arcas, B.A.y., (2017). Communication-Efficient Learning of Deep Networks from Decentralized Data. Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. Fort Lauderdale, USA, 1273-1282.

Önlü, A.Ö., Akca, B., Buyuktanir, B., Yildiz, K. and Baydogmus, G.K., (2025). An investigation and performance evaluation of aggregation algorithms in federated learning architecture. Innovations and Applications in Smart Systems Conference (ASYU). Istanbul, Turkey.

Ozturk, O., Buyuktanir, B., Baydogmus, G.K. and Yildiz, K., 2025. Differential Privacy in Federated Learning: Mitigating Inference Attacks with Randomized Response. arXiv preprint arXiv:2509.13987. https://doi.org/10.48550/arXiv.2509.13987

Paleyes, A., Urma, R.G. and Lawrence, N.D., 2022. Challenges in Deploying Machine Learning: A Survey of Case Studies. *ACM Computing Surveys*, 55, **6**, 114. https://doi.org/10.1145/3533378

Raj, A., Sharma, V. and Shanu, A.K., (2022). Comparative Analysis of Security and Privacy Technique For Federated Learning In IOT Based Devices. 2022 3rd International Conference on Computation, Automation and Knowledge Management (ICCAKM). Dubai, United Arab Emirates, 1-5.

Sasaki, Y., 2022. A Survey on IoT Big Data Analytic Systems: Current and Future. *IEEE Internet of Things Journal*, 9, **2**, 1024-1036. https://doi.org/10.1109/JIOT.2021.3131724

Shah, D., Patel, S. and Bharti, S.K., 2020. Heart Disease Prediction using Machine Learning Techniques. *SN Computer Science*, 1, **6**, 345. https://doi.org/10.1007/s42979-020-00365-y

Sun, Y., Li, H., Shen, Y., Xie, J., Zhao, Y., Gao, X. and Si, N., (2023). An Asynchronous Federated Learning Algorithm Based on a Backup Update of Model Version Parameters. 2023 3rd International Conference on Electrical Engineering and Control Science (IC2ECS). Nanjing, China, 1538-1544.

Surur, F.M., Mamo, A.A., Gebresilassie, B.G., Mekonen, K.A., Golda, A., Behera, R.K. and Kumar, K., 2025. Unlocking the power of machine learning in big data: a scoping survey. *Data Science and Management*, 8, **4**, 519-535. https://doi.org/10.1016/j.dsm.2025.02.004

Tosun, B.E., Yorulmaz, E., Ergul, F.E., Buyuktanir, B., Baydogmus, G.K. and Yildiz, K., (2025). Comparative Analysis of Relational Classification Algorithms on Benchmark Datasets. 2025 10th International Conference on Computer Science and Engineering (UBMK). 790-795.

Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J. and Zhang, W., 2023. A survey on federated learning: challenges and applications. *International Journal of Machine Learning and Cybernetics*, 14, **2**, 513-535. https://doi.org/10.1007/s13042-022-01647-y

Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W. and Gao, Y., 2021. A survey on federated learning. *Knowledge-Based Systems*, 216, 106775. https://doi.org/10.1016/j.knosys.2021.106775

Zhang, T., Zhu, T., Gao, K., Zhou, W. and Yu, P.S., 2023. Balancing Learning Model Privacy, Fairness, and Accuracy With Early Stopping Criteria. *IEEE Transactions on Neural Networks and Learning Systems*, 34, **9**, 5557-5569. https://doi.org/10.1109/TNNLS.2021.3129592

Zhou, S. and Li, G.Y., 2023. FedGiA: An Efficient Hybrid Algorithm for Federated Learning. *IEEE Transactions on Signal Processing*, 71, 1493-1508. https://doi.org/10.1109/TSP.2023.3268845

Internet References

American Heart Association, AHA names biggest advances in cardiovascular research for 2024, https://www.heart.org/en/around-the-aha/aha-names-biggest-advances-in-cardiovascular-research-for-2024, (04.01.2026)

Centers for Disease Control and Prevention (CDC), Diabetes Health Indicators Dataset, http://www.kaggle.com/datasets/alexteboul/diabetes-health-indicators-dataset/data, (04.01.2026)

Fayez, O., Heart Attack Data, https://www.kaggle.com/datasets/omarfayez/heart-attack-data, (04.01.2026)