



A Federated Explainable AI Model for Breast Cancer Classification

Eleni Briola
Department of Electrical and
Computer Engineering, Democritus
University of Thrace
Xanthi, Greece
ebriola@ee.duth.gr

Christos Chrysanthos
Nikolaïdis
Department of Electrical and
Computer Engineering, Democritus
University of Thrace
Xanthi, Greece
cnikolai@ee.duth.gr

Vasileios Perifanis
Department of Electrical and
Computer Engineering, Democritus
University of Thrace
Xanthi, Greece
vperifan@ee.duth.gr

Nikolaos Pavlidis
Department of Electrical and
Computer Engineering, Democritus
University of Thrace
Xanthi, Greece
npavli@ee.duth.gr

Pavlos S. Efraïmidis
Department of Electrical and
Computer Engineering, Democritus
University of Thrace
Xanthi, Greece
pefraimi@ee.duth.gr

ABSTRACT

Breast cancer diagnosis is a crucial domain where Explainable Artificial Intelligence (XAI) integration holds immense importance. Understanding AI model decisions not only enhances trust but also aids in treatment strategies. However, the need for explainability must address privacy concerns, prompting the exploration of Federated Learning. This study explores the intersection of Explainable AI, Privacy, and Federated Learning in breast cancer diagnosis. Utilizing Wisconsin Diagnostic Breast Cancer Dataset and Wisconsin Breast Cancer Dataset, our results showcase that Federated Learning enhances user privacy while maintaining performance, achieving an accuracy of 97.59% and F1 score of 98.393% in Wisconsin Diagnostic Breast Cancer Dataset using artificial neural networks and 97.14% accuracy and 95.65% F1 score in Wisconsin Breast Cancer Dataset employing XGBoost. By computing SHAP values locally, we maintain explainability while enhancing privacy. Our findings highlight the potential of federated learning in maintaining privacy and explainability, advancing breast cancer diagnosis and treatment.

CCS CONCEPTS

• **Computing methodologies** → **Artificial intelligence; Machine learning**; • **Security and privacy**;

KEYWORDS

Federated Learning, Explainable AI, Federated Explainable AI, Neural Networks, Breast Cancer Classification

ACM Reference Format:

Eleni Briola, Christos Chrysanthos Nikolaïdis, Vasileios Perifanis, Nikolaos Pavlidis, and Pavlos S. Efraïmidis. 2024. A Federated Explainable AI Model for Breast Cancer Classification. In *European Interdisciplinary Cybersecurity Conference (EICC 2024)*, June 05–06, 2024, Xanthi, Greece. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3655693.3660255>

1 INTRODUCTION

Breast cancer diagnosis stands as a critical domain where the integration of Explainable Artificial Intelligence (XAI) holds profound significance. The ability to interpret and understand the decisions made by Artificial Intelligence (AI) models in diagnosing breast cancer not only enhances trust and acceptance but also provides valuable insights for medical practitioners to refine treatment strategies.

However, in the need for explainability, concerns regarding data privacy are significant, especially given the sensitive nature of patient data. Safeguarding patient privacy while harnessing the power of AI is essential. Federated Learning (FL) emerges as a promising field, allowing collaborative model training across decentralized data sources without the need to centralize sensitive information.

Moreover, the challenge for explainability itself can be resolved with the needs of privacy through federated approaches. By computing Shapley values locally on models trained on each client's data, explainability becomes achievable without compromising privacy. This approach ensures that sensitive patient data remain within the boundaries of individual institutions or clients, thus mitigating privacy risks, while maintaining the accuracy and precision of diagnostic outcomes.

In this paper, we explore the intersection of Explainable AI, Privacy, and Federated Learning in the context of breast cancer diagnosis. We demonstrate how federated approaches enable the realization of explainability in AI models while maintaining strict privacy standards, ultimately fostering trust and efficacy in AI-assisted medical diagnostics.

The sections of this paper are structured as follows: Section 2 provides an overview of the related work in the field. In Section 3,



This work is licensed under a Creative Commons Attribution International 4.0 License.

EICC 2024, June 05–06, 2024, Xanthi, Greece
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1651-5/24/06
<https://doi.org/10.1145/3655693.3660255>

we introduce the datasets utilized in this study, offering concise descriptions for each dataset. Section 4 elaborates on the methodology adopted for this research effort. Following this, in Section 5, we present the evaluation of experiments conducted. Finally, Section 6 encapsulates the key findings derived from this study and outlines potential avenues for future research.

2 RELATED WORK

2.1 Breast Cancer Detection

Advancements in Machine Learning (ML) and Deep Learning (DL) have significantly impacted the field of medical imaging, particularly in breast cancer detection and diagnosis. This section reviews recent systematic studies and research efforts that leverage these computational techniques to improve the accuracy, efficiency, and clinical applicability of breast cancer screening methods.

Recent research has explored various machine learning methodologies for breast cancer detection, with a notable emphasis on deep learning approaches due to their superior performance in image analysis and pattern recognition. Gardezi et al.[3] conducted a systematic review focusing on the evolution of machine learning methods in mammographic data analysis for breast cancer diagnosis. They highlighted the transition from traditional ML to DL techniques, underscoring the latter's potential in enhancing the diagnostic capabilities of computer-aided diagnosis (CAD) systems. In a comprehensive review, Nasser and Yusof [11] examined the effectiveness of deep learning techniques in breast cancer detection, emphasizing the predominance of Convolutional Neural Networks (CNNs). Their findings suggest that CNNs significantly improve early diagnosis and, consequently, patient survival rates.

Mao et al.[8] investigated the application of machine learning models in ultrasound elastography for breast tumor classification. They reported that while deep learning models, such as CNNs, show promise, they do not necessarily outperform traditional imaging workflows. Furthermore Yu et al.[21] provided a survey on the deployment of deep learning in breast cancer CAD systems, highlighting the rapid development and potential of deep neural networks to surpass human-level performance in certain diagnostic tasks.

The integration of machine learning and deep learning techniques in breast cancer detection and diagnosis represents a significant shift towards more accurate, efficient, and patient-centered healthcare. As these technologies continue to evolve, their application in clinical settings promises to further enhance the early detection and treatment of breast cancer, ultimately improving patient outcomes.

2.2 Explainable AI

In the realm of breast cancer classification using Explainable AI, the work presented by Khater et al. [6] stands as a significant contribution, aiming to develop a model for classifying breast cancer and providing insights into the underlying factors influencing classification decisions. Their study demonstrates the effectiveness of machine learning techniques, achieving high accuracy and precision rates on the Wisconsin breast cancer and diagnostic breast cancer datasets. Notably, their research emphasizes the importance of interpretability and transparency in machine learning models

applied to healthcare, shedding light on crucial features such as bare nuclei and the area's worst feature in diagnosing breast cancer malignancy. Additionally, the use of SHAP (SHapley Additive exPlanations) values, a technique for explaining individual predictions in machine learning models, further enhances the interpretability of their model. SHAP values provide insights into how each feature contributes to the model's decision-making process, offering clinicians valuable information to understand and trust the model's predictions.

Building upon the foundation laid by Khater et al. [6], our work innovatively extends the application of Explainable AI in breast cancer classification by adopting a federated learning approach. By training the model in a federated manner, our research addresses concerns of data privacy and security while maintaining the high performance achieved in centralized settings. This novel approach not only enhances the accessibility and applicability of machine learning models in healthcare but also underscores the significance of advancing interpretability and transparency in artificial intelligence-based medical systems.

2.3 Federated Learning

In the current era, the deployment of deep learning techniques is fragmented by data privacy and security concerns. Regulatory frameworks, such as the General Data Protection Regulation (GDPR), impose stringent requirements on data processing, aiming to protect individual privacy. The inherent reliance of learning algorithms on vast datasets for training, complicates data acquisition and processing due to privacy considerations, often undermining their effectiveness. In response to these challenges, Federated Learning (FL) [10] has emerged as a promising approach. FL is a privacy-centric approach that allows training models in a distributed manner, without requiring users to share their associated data.

The federated process (see Fig. 1) begins with a central authority initializing a model's parameters, acting as the coordinator of the distributed training process. Federated training is organized in discrete time intervals, called federated rounds. Each federated round involves three main steps, that are executed iteratively:

- The coordinator selects a set of available clients using a sampling algorithm and broadcasts them the current global parameters.
- Selected clients perform local training on the received model parameters and transmit back to the coordinator the local model parameters.
- The coordinator server applies an aggregation on the received local models to generate the global parameters for the next round.

FL has been applied across various domains, including cellular networks [14], recommender systems [13], social-care [12] and healthcare [16]. Especially, for the healthcare domain, where sensitive data should be processed, FL has been proven particularly useful as a key-enabler for privacy-preserving machine learning operations.

Jiménez-Sánchez et al. [4] introduced a memory-aware FL approach for breast cancer classification using mammography, effectively handling class imbalance and data heterogeneity across different institutions. Jindal et al. [5] explored the application of

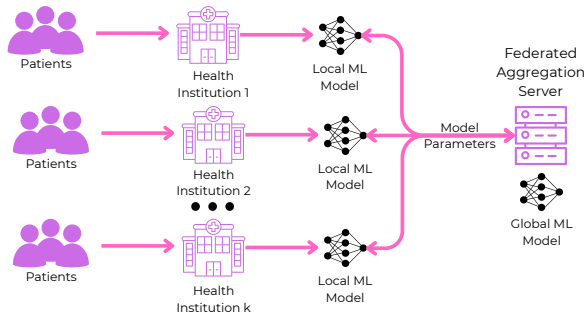


Figure 1: Federated Training Architecture

FL with Convolutional Neural Networks (CNNs) for lung cancer diagnosis, demonstrating the model’s scalability and its adaptability to local data distributions. Repetto and Torre [15] presented a FL framework through goal programming that proves suitable for various cancer detection tasks, emphasizing privacy compliance and local training efficiency. Finally, Maurya et al. [9] applied FL for colorectal cancer prediction, demonstrating results comparable to centralized models.

Previous research in FL has explored its applicability within healthcare, demonstrating the potential to enhance collaborative efforts without compromising patient privacy. However, these efforts often overlook a critical aspect of AI applications, especially in healthcare, i.e., explainability. Understanding the rationale behind a model’s predictions is crucial since clinicians require insights into the decision-making process of models to trust and effectively utilize AI tools for diagnosis and treatment planning. To this end, this work integrates XAI mechanisms within the framework of FL for breast cancer prediction. We consider a scenario where N clinicians, each owning patient-based data, collaborate to enhance predictive accuracy while preserving patient confidentiality. After model training, we employ advanced XAI techniques to elucidate the reasons behind the model’s predictions, aiming to bridge the trust gap between AI models and healthcare professionals.

2.4 Federated Explainable AI

The fusion of Federated Learning (FL) principles with Explainable AI (XAI) methodologies has drawn escalating attention in recent times. Some papers refer to it as fXAI [7], while others refer to it as FEDXAI [1]. In the study [17], Wang suggests the utilization of Shapley values for interpreting Federated Learning models, particularly KNN, without compromising data privacy. The experimental findings demonstrate consistent feature importance results for a subset of features in comparison to results obtained from considering all features. Furthermore, the Shapley value of the combined federated feature offer a meaningful insight into the overall contribution of federated features from the guest party, while keeping guest data entirely concealed from the host’s perspective. Transitioning to broader conceptual frameworks, Bárcena et al. [1] outline the concepts of Federated Learning (FL), Explainable AI (XAI), and Federated Explainable AI (FED-XAI), along with key challenges and

potential solutions. They suggest that while FED-XAI is in its early stages, its ability to ensure data privacy and explainability, while maintaining performance, indicates its potential for widespread adoption in future AI applications.

The emergence of Federated Explainable Artificial Intelligence (fXAI) presents a significant advancement in machine learning methodologies, exemplified by the XRule algorithm proposed by Kusiak [7]. Operating within a collaborative environment similar to federated learning, fXAI focuses on constructing transparent and interpretable models to cater to diverse user needs, extending the concept of explainable AI to distributed decision-making scenarios. Through practical examples and industrial applications, fXAI showcases its potential to address complex challenges, offering a versatile framework seamlessly integrating explicit data science algorithms.

3 DATASETS

In this section, we provide an in-depth overview of both the Wisconsin Breast Cancer Dataset and Wisconsin Diagnostic Breast Cancer Dataset that were employed in our experiments. These datasets are extensively utilized in numerous studies and research initiatives focusing on classification and prediction tasks within this domain. Through a comprehensive examination of the features within these datasets, our aim is to enhance comprehension and recognition of their importance in supporting precise and trustworthy classification and prediction tasks.

3.1 Wisconsin Breast Cancer Dataset

The WBC dataset [18], a widely-used dataset in ML and health analysis, provides information on breast cancer tumors for classification and regression tasks. It comprises details like tumor size, shape, and texture, among others. Introduced in 1992 by Dr. William H. Wolberg [20], it includes 699 instances, each with 10 features as shown in Table 1.

Table 1: Wisconsin Breast Cancer Dataset Description

Parameters	Values
Number of instances	699 instances
Number of features	10 features
Number of inputs	9 inputs (each input has from a scale 1: 10)
Number of outputs	two outputs (Malignant and Benign)

The first nine features describe tumor characteristics, while the last indicates malignancy. Each feature is evaluated on a scale from 1 to 10, with a score of 1 indicating characteristics closer to benign traits and a score of 10 indicating characteristics closer to malignant traits[20].

3.2 Wisconsin Diagnostic Breast Cancer Dataset

The WDBC dataset is a publicly available dataset [19] comprising medical records from breast cancer patients. Dr. William H. Wolberg from the University of Wisconsin Hospitals collected this dataset in the early 1990s, and it serves as a widely utilized resource for

Table 2: Wisconsin Diagnostic Breast Cancer Dataset Description

Parameters	Values
Number of instances	568 instances
Number of features	30 features
Number of inputs	28 inputs (the inputs vary from 0 to 2500)
Number of outputs	two outputs (Malignant and Benign)

ML algorithm research and development. With 569 observations, each entry contains 30 attributes as shown in Table 2.

The initial attribute is a unique ID number assigned to each patient, followed by the diagnosis of breast cancer, categorized as either malignant or benign. The remaining 28 attributes detail various tumor characteristics, encompassing size, shape, and texture.

4 METHODOLOGY

This section outlines the methodologies utilized in this research to create a framework for federated explainable machine learning for diagnosing breast cancer. The procedures employed were as follows:

4.1 Data Collection and Preprocessing

The initial phase of this study involves gathering the Wisconsin Breast Cancer (WBC) dataset and the Wisconsin Diagnostic Breast Cancer Dataset (WDBC), which contain comprehensive information about women diagnosed with breast cancer.

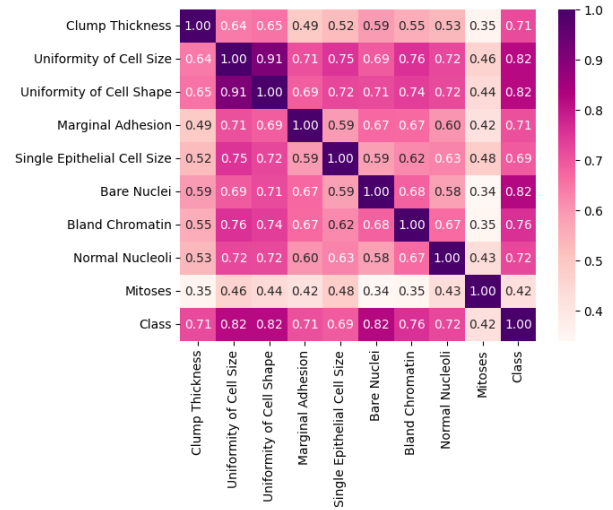
Upon collecting the data, the subsequent step involves cleaning and preprocessing. In the WBC dataset this includes scaling the data to ensure uniformity across features and removing any duplicate values. Furthermore, the class labels indicating tumor diagnosis are transformed from 2 (benign tumors) and 4 (malignant tumors) to 0 and 1, respectively. In the WDBC dataset the diagnosis column containing 'M' (malignant) and 'B' (benign) labels is converted to numerical values, where 'M' is replaced with 1 and 'B' with 0. Certain features are transformed to address skewness and scale issues. Specifically, logarithmic transformations are applied to features such as 'radius_se', 'perimeter_se', 'area_se', 'concavity_se', and 'fractal_dimension_se'. Outliers in the dataset are detected using Z-score analysis, and rows containing outliers are filtered out from the dataset. Standardization is performed on the selected features using `StandardScaler()` from scikit-learn Python Library to ensure that all features have a mean of 0 and a standard deviation of 1. The standard score of a sample x is calculated as:

$$z = \frac{x - u}{s} \quad (1)$$

where u is the mean of the training samples, and s is the standard deviation of the training samples.

4.2 Feature Identification, Algorithm Selection, and Model Training

Following data cleaning and preprocessing, significant features for the model are identified using feature importance techniques provided by machine learning algorithms. For the WBC Dataset, it is

**Figure 2: Feature correlation matrix for WBC dataset**

concluded from the correlation matrix in Figure 2 that the uniformity of cell size and uniformity of cell shape are highly correlated. Additionally, these features also exhibit a high correlation with the output class as well as the bare nuclei feature. Therefore, the "Uniformity of Cell Shape" was removed. For the WDBC Dataset, only the "worst" characteristics were selected. The correlation matrix of "worst" characteristics is shown in Figure 3. Firstly, dimensionality reduction is achieved by selecting 10 out of 30 features. Secondly, the ML model demonstrates higher predictive power. Compared to the other features in the dataset, the "worst" features exhibit a stronger correlation with the prevalence of malignancy. Concentrating solely on these characteristics may enhance the accuracy of a model in identifying whether a tissue sample is benign or malignant. Finally, fewer features in a model can improve explainability, making it simpler to analyze and comprehend the variables influencing the model's predictions.

Once the relevant features are determined, an appropriate machine-learning algorithm is chosen to address the breast cancer diagnosis problem. For the WBC dataset, these algorithms are Support Vector Machine (SVM), Random Forest, XGBoost, and k-Nearest Neighbors (KNN). For the WDBC Dataset, these algorithms are Artificial Neural Network (ANN), Random Forest, XGBoost, and Support Vector Machine (SVM). We utilized the same algorithms as Khater et al. [6] in order to be able to compare our results.

The selected algorithms are trained on the preprocessed dataset using a portion of the data for training (75%) and another portion for testing (25%) to evaluate the model's performance. Following training, the model's performance is assessed using the testing dataset. Evaluation metrics such as accuracy, precision, and F1 score are calculated to see the model's diagnostic capabilities.

Subsequently, the model's output and predictions are analyzed and interpreted using Explainable AI (XAI) techniques. Specifically, permutation importance and Shapley values are employed to gain insights into the model's decision-making process.

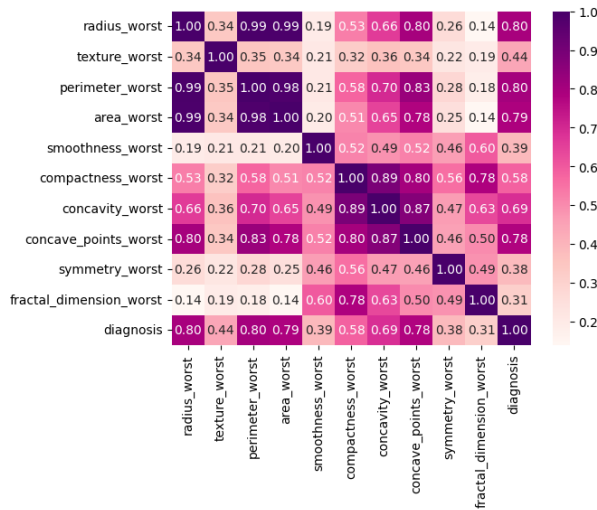


Figure 3: Feature correlation matrix for WDBC dataset

4.3 Federated Learning Approach

The methodology is replicated for federated learning, where data privacy is a concern. In this setup, Explainable AI (XAI) techniques are applied to the federated global model to ensure transparency and explainability while maintaining privacy. For federated learning, we utilized the Flower framework. The Flower framework has been presented by Beutel et al. [2] is an open-source platform designed to simplify Federated Learning experiments, thus making it accessible to a wider audience. Flower streamlines the process of conducting Federated Learning experiments by providing a user-friendly codebase, built-in communication protocols, and compatibility with popular machine learning frameworks such as PyTorch and TensorFlow. For the WBC dataset, Federated XGBoost was employed since XGBoost exhibited the best accuracy in centralized training. For the WDBC dataset, Federated Artificial Neural Network (ANN) was utilized as ANN demonstrated the best performance in centralized training.

For the federated XGBoost, the following steps were followed: Firstly, a utility function for XGBoost trees was defined to facilitate the federated learning process. Subsequently, a tabular dataset was partitioned into four clients evenly to prepare the data for federated learning. Global variables necessary for federated XGBoost learning were defined to maintain consistency across clients. Next, a global XGBoost tree was built for comparison purposes. Local XGBoost trees were then simulated on clients for further comparison. The process of centralized federated XGBoost was executed to ensure proper coordination. Custom client and server components were created using the Flower framework to facilitate communication and coordination among clients and the server. Server-side evaluation and experimentation were conducted to assess the performance of the federated XGBoost model. Finally, SHAP was applied to interpret the predictions of the federated global model and the predictions of each federated client model, enhancing transparency and explainability. In order to be able to compare the results of the federated global model and each client's model, we aggregated

the shap values of the local models using an aggregated weighted average.

For the federated ANN, the following steps were followed: Initially, centralized training with PyTorch was conducted using a neural network architecture comprising three layers with respective neuron counts of 15, 10, and 1, and the number of epochs set to 100. Subsequently, a Flower client was defined to facilitate federated learning. The dataset was then split equally among three clients randomly. A Flower strategy was chosen to aggregate federated metrics and a callback function was defined to evaluate the state of the global model on a centralized dataset. Finally, Shapley Values were calculated using SHAP Python Library to interpret the predictions of the federated global model and each client's model as in XGBoost.

5 EXPERIMENTAL EVALUATION

In this section, we present and discuss the outcomes obtained from employing Federated Learning to classify breast cancer. The evaluation is conducted on two distinct datasets: WBC and WDBC.

5.1 Evaluation of WBC Dataset

The study employed various machine learning (ML) algorithms including SVM, KNN, RF, and XGBoost. Training the ML model centralized with these algorithms revealed that XGBoost achieved the highest accuracy, as illustrated in Table 3. Consequently, XGBoost was selected for FL. Following the assessment of the ML model's performance, it becomes crucial to analyze the findings to comprehend its behavior. To achieve this, two model-agnostic techniques, permutation importance, and Shapley Values were utilized.

Table 3: Evaluation Results of WBC Dataset

	Accuracy	F1	Precision
SVM	95.32%	93.84%	98.38%
Random Forest	97.07%	96.26%	98.46%
XG-Boost	97.14%	95.65%	98.41%
KNN	96.32%	94.84%	98.38%

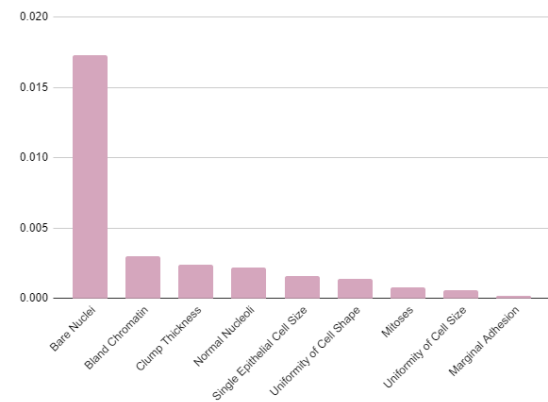


Figure 4: Permutation importance analysis in WBC dataset

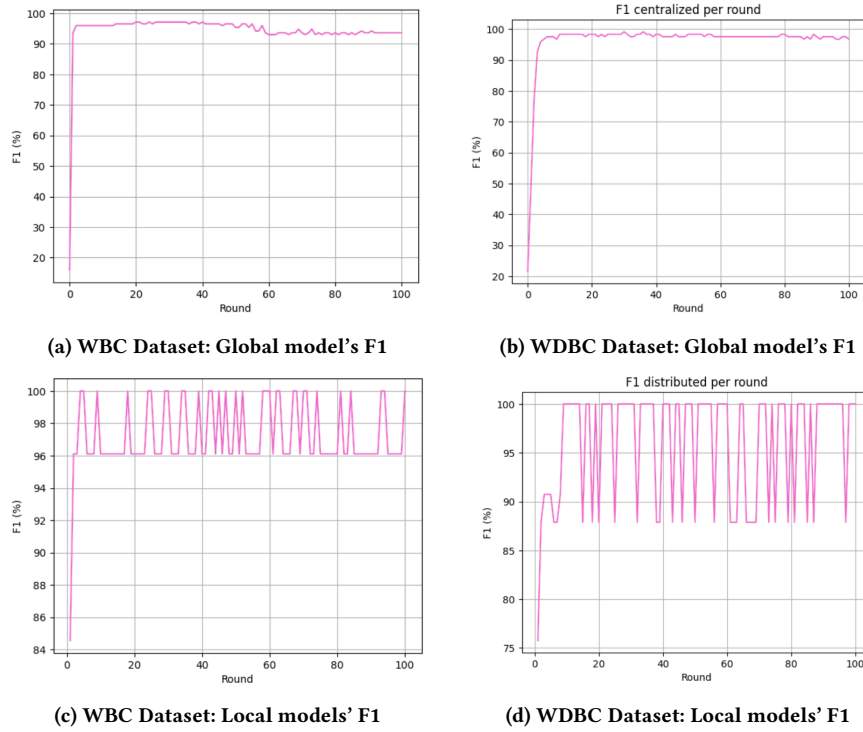


Figure 5: WBC and WDBC datasets Federated Learning Results

This figure presents the results of federated learning on the WBC and WDBC datasets, showing the performance of the federated ANN (5b, 5d) and the federated XG-Boost (5a, 5c)

The permutation importance technique is employed to prioritize features, enabling the identification of the most impactful ones. Illustrated in Figure 4, Bare nuclei emerges as the foremost feature. Following this, Shapley values are employed to assess the relative significance of features in individual predictions. A SHAP summary plot is produced to assess the role of features in breast cancer classification. Figure 6 demonstrates that the Bare nuclei feature maintains the greatest contribution, consistent across both centralized and distributed training, affirming the permutation findings. These findings align with those reported in Khater et al. [6].

During distributed training, where models are trained on local data and aggregated to form a global model, it is noteworthy that both the global model (Figure 6b) and the aggregated local models (Figure 6c) yielded comparable outcomes. This observation underscores the effectiveness of federated learning in maintaining model performance across decentralized data sources, thereby affirming the reliability and robustness of the approach.

5.2 Evaluation of WDBC Dataset

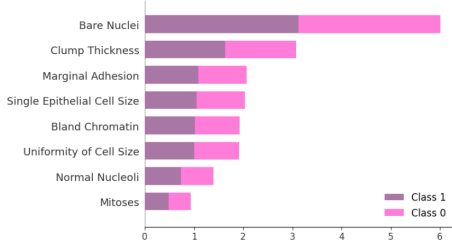
The study employed various machine learning (ML) algorithms including ANN, SVM, RF, and XGBoost. Training the ML model with these algorithms revealed that ANN achieved the highest accuracy, as presented in Table 4. Consequently, ANN was chosen for Federated Learning.

Table 4: Evaluation Results of WDBC Dataset

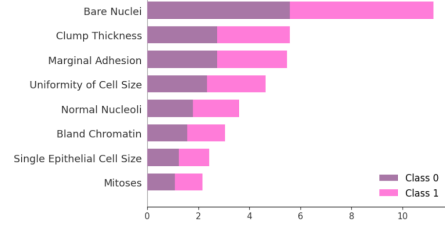
	Accuracy	F1
ANN	98.39%	98.39%
Random Forest	95.96%	94.50%
XG-Boost	97.58%	97.58%
SVM	98.38%	98.38%

In order to develop an interpretable Federated Learning model, Shapley values were calculated for features that represented extreme values to ascertain their influence on breast cancer class prediction. Remarkably, the feature "area worst," signifying the total area occupied by the nucleus, emerged as the most significant factor in breast cancer classification. The findings align with those reported in Khater et al. [6].

The SHAP plot depicted in Figure 7 reveals that increased values of the area feature have a positive impact on the classification task in both centralized and distributed training settings. Specifically, higher area values correlate with elevated predictions of malignant breast cancer. It is important to note the significance of federated learning in this context. During distributed training, noteworthy findings emerged. Both the global model (Figure 7b) and the aggregated local models (Figure 7c) exhibited comparable performance.



(a) Centralized training



(b) Federated training global model

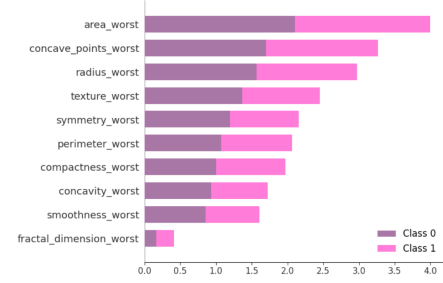
(c) Distributed training averaged Shapley values of local models

Figure 6: Shapley Values for WBC dataset

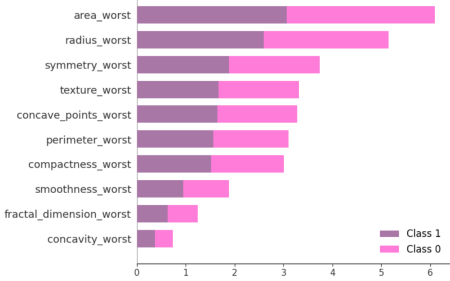
In the centralized training scenario (Figure 7a), the top three influential features for breast cancer classification were identified as area worst, concave points worst, and symmetry worst. Conversely, in the federated setting, both the global and local models ranked area worst, radius worst, and symmetry worst as the most influential features. This divergence in feature importance highlights the adaptability of the federated approach to local data characteristics, which may vary across different clients or institutions.

Furthermore, when considering the least influential features, fractal worst, smoothness worst, and concavity worst emerged as the bottom three in centralized training. In contrast, in the federated setting, both the global and local models ranked concavity worst, fractal worst, and smoothness worst as the least influential features.

Overall, these findings highlight the flexibility and adaptability of federated learning in accommodating diverse data sources while enhancing model explainability and performance. By acknowledging and accommodating such variations in feature importance, federated learning enables robust and reliable model training across decentralized data environments, thereby enhancing the effectiveness and applicability of AI-driven solutions in healthcare and other sensitive domains.



(a) Centralized training



(b) Federated training global model

(c) Distributed training averaged Shapley values of local models

Figure 7: Shapley Values for WDBC dataset

6 CONCLUSION AND FUTURE WORK

In conclusion, this study highlights the transformative potential of Federated Explainable Artificial Intelligence (FEDXAI), particularly in the realm of breast cancer diagnosis. Our research has demonstrated the effectiveness of federated learning models in accurately classifying breast cancer, with notable achievements in accuracy and model explainability. By leveraging Federated Learning (FL), we have addressed privacy concerns associated with sensitive medical data, ensuring user privacy without compromising model performance. Our findings underscore the significance of explainability and transparency in AI-driven healthcare systems, showing the way for improved breast cancer diagnosis.

Moving forward, there are numerous avenues for advancing research and development in this domain. One particularly promising avenue involves exploring diverse methods of aggregating XAI results from local models, potentially leading to more robust and

accurate predictions, thereby enhancing the effectiveness of Federated Learning in breast cancer diagnosis. Furthermore, delving into innovative XAI techniques tailored specifically to the challenges of medical data could offer valuable insights and enhance the performance of AI-driven healthcare systems. Additionally, leveraging encryption techniques such as secure multiparty computations can ensure privacy. Overall, the integration of XAI techniques into Federated Learning holds great promise for future research, offering substantial potential to advance breast cancer diagnosis significantly.

REFERENCES

- [1] José Luis Corcuera Bárcena, Mattia Daole, Pietro Ducange, Francesco Marcelloni, Alessandro Renda, Fabrizio Ruffini, and Alessio Schiavo. 2022. Fed-XAI: Federated Learning of Explainable Artificial Intelligence Models.. In *XAI. it@ AI* IA*. 104–117.
- [2] Daniel J Beutel, Taner Topal, Akhil Mathur, Xinchu Qiu, Javier Fernandez-Marques, Yan Gao, Lorenzo Sani, Hei Li Kwing, Titouan Parcollet, Pedro PB de Gusmão, and Nicholas D Lane. 2020. Flower: A Friendly Federated Learning Research Framework. *arXiv preprint arXiv:2007.14390* (2020).
- [3] S. J. S. Gardezi, A. Elazab, Baiying Lei, and Tianfu Wang. 2019. Breast Cancer Detection and Diagnosis Using Mammographic Data: Systematic Review. *Journal of Medical Internet Research* 21 (2019). <https://doi.org/10.2196/14464>
- [4] Amelia Jiménez-Sánchez, Mickael Tardy, Miguel A González Ballester, Diana Mateus, and Gemma Piella. 2023. Memory-aware curriculum federated learning for breast cancer classification. *Computer Methods and Programs in Biomedicine* 229 (2023), 107318.
- [5] Varun Jindal, Vinay Kukreja, Devesh Pratap Singh, Satvik Vats, and Shiva Mehta. 2023. Modernizing Lung Cancer Detection: The Federated Learning CNN Approach. In *2023 4th IEEE Global Conference for Advancement in Technology (GCAT)*. IEEE, 1–6.
- [6] Tarek Khater, Abir Hussain, Riyad Bendardaf, Iman M. Talaat, Hissam Tawfik, Sam Ansari, and Soliman Mahmoud. 2023. An Explainable Artificial Intelligence Model for the Classification of Breast Cancer. *IEEE Access* (2023), 1–1. <https://doi.org/10.1109/ACCESS.2023.3308446>
- [7] Andrew Kusiak. 2024. Federated explainable artificial intelligence (fXAI): a digital manufacturing perspective. *International Journal of Production Research* 62, 1-2 (2024), 171–182. <https://doi.org/10.1080/00207543.2023.2238083>
- [8] Ye-Jiao Mao, Hyo-Jung Lim, Ming Ni, Wai-Hin Yan, D. Wong, and James Siu Wai Cheung. 2022. Breast Tumour Classification Using Ultrasound Elastography with Machine Learning: A Systematic Scoping Review. *Cancers* 14 (2022). <https://doi.org/10.3390/cancers14020367>
- [9] Yash Maurya, Prahaladh Chandrahasan, and G Poornalatha. 2022. Federated Learning for Colorectal Cancer Prediction. In *2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT)*. IEEE, 1–5.
- [10] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. 2017. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*. PMLR, 1273–1282.
- [11] Maged Nasser and U. K. Yusof. 2023. Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction. *Diagnostics* 13 (2023). <https://doi.org/10.3390/diagnostics13010161>
- [12] Christos Chrysanthos Nikolaidis, Vasileios Perifanis, Nikolaos Pavlidis, and Pavlos S Efraimidis. 2023. Federated Learning for Early Dropout Prediction on Healthy Ageing Applications. In *2023 Eighth International Conference on Fog and Mobile Edge Computing (FMEC)*. IEEE, 135–142.
- [13] Vasileios Perifanis and Pavlos S Efraimidis. 2022. Federated neural collaborative filtering. *Knowledge-Based Systems* 242 (2022), 108441.
- [14] Vasileios Perifanis, Nikolaos Pavlidis, Remous-Aris Koutsiamanis, and Pavlos S Efraimidis. 2023. Federated learning for 5G base station traffic forecasting. *Computer Networks* 235 (2023), 109950.
- [15] Marco Repetto and Davide La Torre. 2022. Federated Learning through Goal Programming: a Computational Study in Cancer Detection. In *2022 5th International Conference on Signal Processing and Information Security (ICSPIS)*. IEEE, 80–85.
- [16] Nicola Rieke, Jonny Hancox, Wenqi Li, Fausto Milletari, Holger R Roth, Shadi Albarqouni, Spyridon Bakas, Mathieu N Galtier, Bennett A Landman, Klaus Maier-Hein, et al. 2020. The future of digital health with federated learning. *NPJ digital medicine* 3, 1 (2020), 119.
- [17] Guan Wang. 2019. Interpret federated learning with shapley values. *arXiv preprint arXiv:1905.04519* (2019).
- [18] William Wolberg. 1992. Breast Cancer Wisconsin (Original). UCI Machine Learning Repository. DOI: <https://doi.org/10.24432/C5HP4Z>.
- [19] William Wolberg, Olvi Mangasarian, Nick Street, and W. Street. 1995. Breast Cancer Wisconsin (Diagnostic). UCI Machine Learning Repository. DOI: <https://doi.org/10.24432/C5DW2B>.
- [20] W H Wolberg and O L Mangasarian. 1990. Multisurface method of pattern separation for medical diagnosis applied to breast cytology. *Proceedings of the National Academy of Sciences* 87, 23 (1990), 9193–9196. <https://doi.org/10.1073/pnas.87.23.9193>
- [21] Xiang Yu, Qinghua Zhou, Shuihua Wang, and Yudong Zhang. 2021. A systematic survey of deep learning in breast cancer. *International Journal of Intelligent Systems* 37 (2021), 152 – 216. <https://doi.org/10.1002/int.22622>