Int. J. Appl. Math. Comput. Sci., 2026, Vol. 36, No. 2, 211–221
DOI: 10.61822/amcs-2026-0015

MANAGING UNCERTAINTY IN FEDERATED LEARNING VIA INTERVAL FUZZY SETS AND ENTROPY–BASED FUSION

JAROSŁAW SZKOŁA [a], BARBARA PĘKALA [a,b,*], KRZYSZTOF DYCZKOWSKI [c]

[a]Faculty of Exact and Technical Sciences
University of Rzeszów
al. Rejtana 16c, 35-959 Rzeszów, Poland
e-mail: jszkola@ur.edu.pl

[b]Department of Artificial Intelligence
University of Information Technology and Management
ul. Sucharskiego 2, 35-225 Rzeszów, Poland
e-mail: bpekala@ur.edu.pl

[c]Faculty of Mathematics and Computer Science
Adam Mickiewicz University
ul. Wieniawskiego 1, 61-712 Poznań, Poland
e-mail: chris@amu.edu.pl

This paper introduces a federated learning framework designed to improve the reliability of diagnostic models under conditions of uncertainty, with a particular focus on medical applications such as breast cancer diagnosis. The proposed method integrates interval-valued fuzzy sets to capture data imprecision and employs logistic regression enhanced with interval-based parameter estimation. Model parameters are aggregated across clients using the Choquet integral, extended with an entropy-based weighting scheme that accounts for both model performance and uncertainty. Experimental results on the Wisconsin breast cancer dataset demonstrate that the proposed federated architecture achieves superior performance compared to traditional methods, particularly in non-IID and unbalanced data scenarios. The framework offers robust privacy preservation, effective uncertainty modeling, and improved classification accuracy, making it suitable for high-stakes, privacy-sensitive domains.

Keywords: federated learning, interval-valued fuzzy sets, uncertainty modeling, Choquet integral, entropy-based aggregation, interval logistic regression, medical diagnostics, non-IID data.

1. Introduction and motivation

Federated learning (FL) has emerged as a promising paradigm in machine learning, particularly in scenarios requiring collaborative model training without direct data sharing. Its core motivation lies in enabling access to distributed datasets while preserving the privacy of the contributing entities. This is especially critical in sensitive domains such as healthcare, where data is often decentralized, privacy-protected, and subject to significant variability and uncertainty.

In this paper, we propose an enhanced federated learning framework that integrates two complementary mechanisms to address uncertainty in learning: (1) an interval-valued approach for training local models and (2) a novel data aggregation method based on the Choquet integral augmented with an entropy measure. These components are specifically designed to handle challenges posed by non-independent and identically distributed (non-IID) as well as unbalanced data distributions, which are common in real-world applications, particularly in medical diagnostics such as breast cancer classification.

Our research focuses on scenarios involving noisy, incomplete, or imbalanced datasets. We investigate how federated learning, when combined

*Corresponding author

with interval-based modeling and entropy-aware aggregation, can improve diagnostic accuracy while maintaining data privacy. The Choquet integral, originally introduced by Choquet (1953), serves as the core aggregation mechanism in our approach. This operator is well-suited for combining interdependent sources of information under uncertainty, and its effectiveness in decision-making has been demonstrated in multiple studies (Szkoła *et al.*, 2024; Pękala *et al.*, 2024; 2025).

A central contribution of this work is the incorporation of interval-valued fuzzy sets (IVFSs) into the learning process. First introduced by Zadeh (1975) and Sambuc (1975), IVFSs extend traditional fuzzy sets by using intervals instead of precise membership degrees. Their mathematical foundation in interval arithmetic allows them to model imprecision more flexibly, making them particularly useful in applications involving uncertain or incomplete data. IVFSs have been successfully applied in domains such as pattern recognition and medical decision support (Bustince *et al.*, 2016).

We propose a federated classification system that combines two key elements: a novel learning mechanism based on data-driven subintervals for local model training and a parameter aggregation scheme using the Choquet integral enhanced with an entropy-based weighting function. The motivation for employing interval weights stems from the need to represent different ranges of input values with varying importance a limitation in standard learning models which typically assign scalar weights to all input values equally. In medical contexts, this distinction is essential; for example, extreme values in blood pressure or temperature often signal pathological conditions, whereas normal ranges may carry less diagnostic weight. Our approach addresses this by learning parameter values specific to subintervals derived from observed training data distributions.

Federated learning enables decentralized knowledge sharing among independent entities while ensuring that raw data remains local. Its practical viability has been demonstrated in applications such as mobile keyboard prediction (Hard *et al.*, 2018) and multi-center healthcare systems (Deist *et al.*, 2017; Zheng *et al.*, 2020). Despite these successes, further methodological improvements are needed to enhance the robustness and generalizability of federated learning, particularly in the presence of uncertainty, data heterogeneity, and privacy constraints. The techniques introduced in this study contribute toward meeting these challenges and broadening the applicability of FL in high-stakes domains.

The methodology presented in this work addresses the pervasive issue of uncertainty in medical data through a bidirectional approach, yielding the following key contributions:

1. a federated learning architecture capable of managing uncertain and/or unbalanced datasets using interval-valued fuzzy sets, with an innovative parameter aggregation strategy based on the Choquet integral enriched by entropy-based weighting;

2. a novel training method for local models that generates interval-valued weights, improving the expressiveness and adaptability of the model and enhancing its diagnostic performance.

1.1. Federated learning. Federated learning (FL) is a decentralized machine learning paradigm that enables multiple parties or institutions to collaboratively train models without sharing raw data (Kairouz *et al.*, 2021). This ensures data privacy and security, as sensitive information remains on local devices or servers. The concept was first introduced by researchers at Google (Konecný *et al.*, 2016; 2017; McMahan *et al.*, 2017), who demonstrated that it is possible to train accurate models across distributed data sources without centralizing the data itself. Since then, FL has gained significant attention and has been applied in various domains, including healthcare, finance, and mobile applications (Yang *et al.*, 2019; Wilbik and Grefen, 2021; Yan *et al.*, 2021).

Unlike traditional distributed learning approaches, which often require centralized data storage and processing, federated learning introduces several distinctive challenges. These include limited communication bandwidth between clients and servers, data heterogeneity (due to non-IID and unbalanced distributions across clients), and the need to maintain strict privacy constraints (Yang *et al.*, 2019; Li *et al.*, 2020). In the FL framework, each participating client trains a copy of the global model on its local dataset, comprising n_k observations, over a predefined number of iterations. The client then transmits only the learned model parameters (or updates) to a central server, rather than the raw data. These updates are subsequently aggregated to form a new global model, which is redistributed to clients in the next training round.

This privacy-preserving model training process enables institutions to leverage diverse and decentralized data sources without compromising confidentiality, a property particularly important in domains such as medical diagnostics and financial systems.

A crucial component of federated learning is the aggregation of model parameters obtained independently by each client. The effectiveness of this aggregation step significantly influences the overall performance of the federated system. An early approach employed a simple arithmetic mean during aggregation. However, subsequent research (Wilbik *et al.*, 2023; 2022) has emphasized the limitations of this method and proposed more sophisticated aggregation schemes, particularly those that incorporate model quality metrics into the weighting process. Notably, methods based on fuzzy integrals, such as the Sugeno and Choquet integrals, have demonstrated superior performance in handling uncertainty and variability across client models (Wilbik *et al.*, 2022; Pękala *et al.*, 2024).

Our study builds upon this line of work by implementing a federated learning process structured around iterative communication and model refinement, as outlined in Fig. 1. The learning protocol follows these steps:

1. Each client begins by training a local model on its private data. After several training epochs, the client sends two types of information to the server: (1) updated model parameters and (2) performance metrics such as accuracy, sensitivity, specificity, precision, and the area under the ROC curve (AUC). These metrics serve as the basis for entropy-weighted aggregation using the Choquet integral.

2. Upon receiving this information, the server aggregates the individual model updates to produce a new global model (line 2).

3. The updated global model is then redistributed to all clients (line 3).

4. Each client integrates the global model into its local learning process, provided the update leads to improved performance (line 4).

5. The aggregation mechanism assigns weights to client models based on both their predictive effectiveness and an entropy-based measure of data uncertainty. These weights are then applied using the Choquet integral during parameter fusion (line 6). The aggregation mechanism assigns weights to each client based on predictive effectiveness and uncertainty. Specifically, the weight w_i for client i is calculated as

$$w_i = \frac{(1 - E_i)Q_i}{\sum_{k=1}^{r}(1 - E_k)Q_k}, \tag{1}$$

where Q_i denotes the client's model quality (AUC) and E_i is the normalized entropy representing model uncertainty. These weights are then applied within the Choquet integral:

$$\rho_i^{agg} = C_g(h), \tag{2}$$

allowing entropy-aware parameter fusion across heterogeneous models.

We particularly aggregated parameters of each local model and each attribute, i.e. ρ_k (see (3)) (after the updates described in Section 3.1), and with respect to uncertainty included in data, i.e., the calculated entropy builds the Choquet integral described in Section 3.2.2.

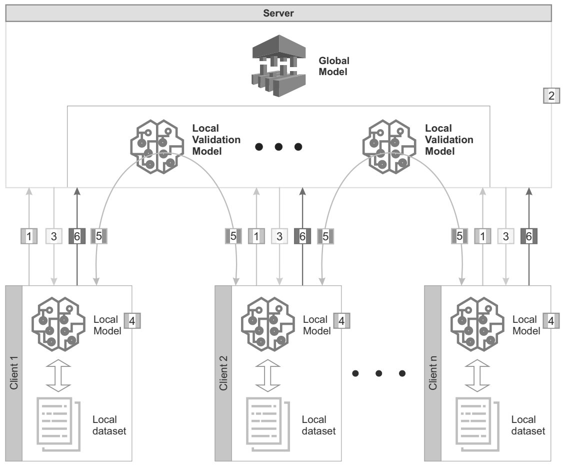

Fig. 1. Extended federated learning model.

This process involves two main stages:

1. In the initial validation phase, the server shares updated models with the clients for evaluation. Each client uses local data to assess whether the updated model provides improved predictive quality.

2. In the second phase, the server collects these evaluations to compute entropy-based weights and aggregates the local models accordingly.

This iterative learning and aggregation process continues until model performance converges, i.e., until subsequent iterations no longer yield significant reductions in the loss function.

A key enhancement of the framework is the inclusion of a cross-validation step for local models. This mechanism allows each client to evaluate its model using external validation sets, without exchanging raw data, thereby reducing the risk of overfitting and improving the generalizability of the resulting global model.

2. Problem and challenges

This study is motivated by key challenges in the field of medical diagnostics, with a particular focus on breast cancer detection. Two major issues underpin our research:

- *Data privacy.* Medical data is highly sensitive and subject to strict regulatory requirements. Ensuring data confidentiality during storage and transfer is essential, particularly in multi-institutional settings.

- *Imprecision and uncertainty in medical data.* Diagnostic data may be incomplete, noisy, or subject to human error and subjective interpretation. These factors arise from limitations in measurement devices and variability in expert assessments.

To effectively address these challenges, our system incorporates the following essential components:

- *uncertainty modeling*, using interval-valued fuzzy sets (IVFSs), with uncertainty quantified in terms of entropy;
- *compatibility with existing diagnostic systems*, allowing integration into current medical infrastructures;
- *robust model validation*, ensuring generalization and reliability of the learned models.

By accounting for both privacy and imprecision, our federated learning system provides a more resilient and interpretable tool for supporting clinical decision-making in uncertain environments.

Structure of the dataset. In our approach, data is represented in the form of intervals to capture inherent uncertainty. Let $L^I = \{[\underline{p}, \overline{p}] : \underline{p}, \overline{p} \in [0,1],\ \underline{p} \leq \overline{p}\}$ denote the set of all closed subintervals within the unit interval. An *interval-valued fuzzy set* (IVFS) (Zadeh, 1975; Sambuc, 1975; Türksen, 1986; Gorzałczany, 1987) on a universe X is a mapping $S : X \to L^I$, where $S(x) = [\underline{S}(x), \overline{S}(x)]$ represents the degree of membership of the element $x \in X$ to the set S. The collection of all such sets is denoted by $\mathrm{IVFS}(X)$.

One of the strengths of the proposed federated learning framework is its model-agnostic nature. For our experiments, we adopt logistic regression due to the binary classification task and the nature of medical data, which typically involves dichotomous outcomes (e.g., malignant vs. benign).

The performance of our uncertainty-aware learning method is evaluated using the Wisconsin Diagnostic Breast Cancer (WDBC) dataset from the UCI Machine Learning Repository (UCI, 2017). This dataset comprises digitized images of fine needle aspirates (FNAs) of breast masses, with numerical features characterizing the morphology of cell nuclei.

For each nucleus, ten real-valued features are measured. To model uncertainty, we construct intervals for each feature by combining the mean and standard deviation of the measurements:

$$[\text{mean} - \text{standard deviation},\ \text{mean} + \text{standard deviation}].$$

Each such interval is then normalized and fuzzified, enabling a robust interval-based representation of the data.

The binary target variable indicates the diagnostic outcome: "0" for malignant and "1" for benign cases. The dataset includes 569 patient samples, with 212 classified as malignant and 357 as benign.

3. Proposed methodology

We propose a federated classification framework that integrates two key innovations: (1) a novel method for training local models using parameterized sub-intervals and (2) an aggregation strategy based on the Choquet integral, incorporating entropy as a measure of uncertainty in local models.

Building upon our previous research (Pękala *et al.*, 2024; 2025), we extend the federated learning framework by explicitly modeling uncertainty through interval-valued weights in the local training phase. This enhancement results in a more expressive and robust learning system that is better suited to handle noisy and heterogeneous data distributions.

As the base learning model, we adopt logistic regression with stochastic gradient descent due to its interpretability, low computational cost, and effectiveness in binary classification tasks. To support interval-based input, we adapt the model to operate on interval-valued features. Specifically, we define the training dataset as $\{Y_i, x_{i1}, \ldots, x_{ip}\}$, where each $x_{ij} \in L^I$ is an interval-valued feature, and $Y_i \in \{0,1\}$ is a binary label for instance $i = 1, \ldots, n$. Here, n denotes the number of instances and p the number of attributes.

During training, each client processes a local dataset of n_k observations through a predefined number of internal iterations. The resulting model learns parameters $\rho = (\rho_0, \rho_1, \ldots, \rho_p)$ and error terms ϵ, and the predictive function for instance i takes the form

$$y_i = \rho_0 + \rho_1 x_{i1} + \cdots + \rho_p x_{ip} + \epsilon_i \quad \text{for} \quad i = 1, \ldots, n_k, \tag{3}$$

where each $\rho_k \in \mathbb{R}$ for $k = 1, \ldots, p$.

Logistic regression is a widely used method across disciplines such as machine learning, healthcare, and the social sciences. It models the relationship between input features and a binary outcome by linearly combining explanatory variables with fixed regression coefficients.

Traditionally, learning consists of adjusting scalar weights ρ_k to minimize the discrepancy between predicted and actual labels. However, this scalar approach fails to account for variations in the influence of input values across different regions of the input domain.

To overcome this limitation, our method introduces a flexible encoding strategy wherein each coefficient is represented not by a single real value, but by a collection of intervals, each associated with a specific subrange of the input feature space. These interval weights are updated dynamically during training, enabling the model to better adapt to local patterns in the data.

This structure allows a single parameter ρ_j to take on multiple scalar values, each corresponding to input values falling within a particular subinterval. As a result, the model can capture non-linear and heterogeneous relationships more effectively than standard logistic regression.

This interval-based parameterization is particularly beneficial in federated settings, where local data distributions may vary significantly between clients, and uncertainty is inherent due to noise, sparsity, or measurement variability.

3.1. Procedure based on interval weights and an entropy measure. Each iteration of the local learning process follows the methodology introduced by Szkoła *et al.* (2024) and is composed of three major stages:

1. initialization of model parameters ρ using sequences of data-dependent sub-intervals;

2. optional optimization of interval widths to merge overly narrow segments;

3. learning procedure: updating interval representatives and weights based on prediction errors.

Procedure 1. Interval weighting and entropy fusion method (IWEF).

Step 1: Initialization of model parameters. Each attribute is associated with a parameter ρ_j, constructed as a sequence of sub-intervals $\rho_{jt} \in L^I$, each linked to a representative value $v_t \in \mathbb{R}$. The complete representation is given by

$$\rho_j = \left\{\rho_{jt}^{\uparrow v_t}\right\} = \left\{[\eta_1, \eta_2)^{\uparrow v_1}, \ldots, [\eta_z, \eta_{z+1}]^{\uparrow v_z}\right\}, \quad 1 \leq j \leq p,$$

where z denotes the number of sub-intervals and p the number of attributes. The intercept term is initialized as $\rho_0 = [0, 1]$ with $v_0 = 1$.

Sub-intervals can be created using one of three strategies: (1) random initialization, (2) fixed-width division, or (3) data-driven segmentation. The third option offers adaptivity by analyzing data distributions to determine boundaries. The process can be refined using a threshold $\psi \in (0, 1)$, which controls whether narrow intervals should be merged.

Step 2: Optimization of interval widths. Narrow intervals can be merged using Algorithm 1, which ensures that all resulting sub-intervals exceed a user-defined significance threshold ψ.

The aggregation function G merges representative values of two intervals using their corresponding weights w_{a_1}, w_{a_2}:

$$G(a_1, a_2) = \begin{cases} \frac{1}{2}\left(\frac{w_{a_1}}{w_{a_2}} a_1 + a_2\right), & \text{if } w_{a_2} \geq w_{a_1}, \\ \frac{1}{2}\left(a_1 + \frac{w_{a_2}}{w_{a_1}} a_2\right), & \text{otherwise.} \end{cases} \tag{4}$$

Algorithm 1: Modification of parameters.

Input: n data instances; p attributes; $\rho_j = \{[\eta_1, \eta_2], \ldots, [\eta_z, \eta_{z+1}]\}$ with representatives $v_1, \ldots, v_z$ for each $j = 1, \ldots, p$; aggregation function G defined in Eqn. (4)

Output: Optimized parameters ρ_j with merged intervals of width $\leq \psi$

```
for j ← 1 to p do
    for k ← 1 to z do
        if η_k − η_{k+1} ≤ ψ then
            if |v_k − v_{k−1}| ≤ |v_k − v_{k+1}| then
                d_i ← G(v_k, v_{k−1}) ← [δ_{k−1}, δ_k]
            else
                d_j ← G(v_k, v_{k+1}) ← [δ_k, δ_{k+1}]
            end
        end
    end
end
```

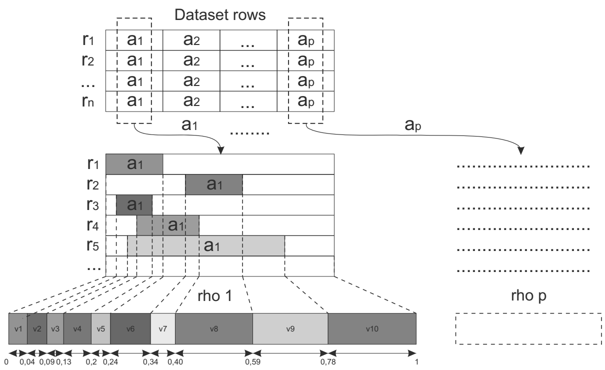

Fig. 2. Initialization of parameter ρ with interval-valued sub-ranges.

Step 3: Learning procedure. Each training iteration consists of the following steps:

1. *Prediction.* For each sample, compute the model output using a logistic sigmoid function:

$$f(y_i) = \frac{1}{1 + e^{-\mathrm{Rep}_\gamma(\rho_0 + \rho_1 x_{i1} + \cdots + \rho_p x_{ip} + \epsilon_i)}},$$

where $\mathrm{Rep}_\gamma(a) = \underline{a} + \gamma_a(\overline{a} - \underline{a})$ is a representative transformation from interval to scalar (see Piegat and Landowski, 2013), with $a = [\underline{a}, \overline{a}]$ and $\gamma_a \in [0, 1]$.

2. *Loss computation.* Using binary cross-entropy loss,

derive

$$\mathcal{L}(y_i) = -\log(f(y_i)) \cdot Y_i - \log(1 - f(y_i)) \cdot (1 - Y_i).$$

3. *Parameter updates.* Update interval-based weights using a gradient descent:

$$\rho_{jt}^{\uparrow v_j} \leftarrow \rho_{jt}^{\uparrow v_j} + \alpha \cdot \nabla_{\rho_j^{\uparrow v_j}} \mathcal{L}(y_i) \cdot x_{ij},$$

$$\rho_0^{\uparrow v_0} \leftarrow \rho_0^{\uparrow v_0} + \alpha \cdot \nabla_{\rho_0^{\uparrow v_0}} \mathcal{L}(y_i),$$

where α is the learning rate and gradients are computed using the representative values of overlapping sub-intervals.

To determine which representatives should be updated, we identify all sub-intervals ρ_j that overlap with the input interval $x_{ij} \in L^I$. If there exists a non-empty intersection with $[\eta_\Omega, \eta_{\Omega+l}]$, then the updated representative is computed as

$$v_j = \sum_{t=\Omega}^{\Omega+l-1} \int_{t+\theta_1}^{t+1-\theta_2} g(t) \cdot v_j,$$

where $g(t)$ is a normalized Gaussian kernel over $[t, t+1]$ and θ_1, θ_2 are offsets controlling overlap sensitivity.

At the end of the learning process, we obtain optimized interval parameters $\rho_{jm}^{\uparrow v_j}$ and $\rho_0^{\uparrow v_0}$ for $1 \leq j \leq p$ and $1 \leq m \leq z$.

Remark 1. During the testing phase (conducted on a randomly selected 10% portion of the dataset), Step 1 is omitted to avoid retraining or re-initialization of interval structures.

3.2. Uncertainty aspect. In this work, we propose an uncertainty-aware aggregation strategy in federated learning by applying the Choquet integral in a novel form enriched with entropy measures, as introduced by Pękala *et al.* (2025). This extension allows us to more effectively incorporate imprecision and heterogeneity of local model parameters, which is especially critical in domains such as medical diagnostics.

To set the stage, we recall fundamental concepts from fuzzy measure theory that underpin the Choquet integral and its application in this context.

Definition 1. Let Z be a finite set. A function $m : 2^Z \to [0, 1]$ is called a *fuzzy measure* (or a monotone measure) if it satisfies the following conditions:

1. $m(\varnothing) = 0$;
2. $m(Z) = 1$;
3. if $A, B \subseteq Z$ and $A \subseteq B$, then $m(A) \leq m(B)$.

A notable subclass of fuzzy measures is the *Sugeno measure* (or the λ-measure), defined by the following recursive condition for any $A, B \subseteq Z$ with $A \cap B \neq \emptyset$:

$$m(A \cup B) = m(A) + m(B) + \lambda \cdot g(A)g(B), \quad (5)$$

where $\lambda > -1$ is a parameter controlling the degree of interaction between sets. For a discrete universe $Z = \{z_1, \ldots, z_n\}$ with $g^j = g(\{z_j\})$, the value of λ satisfies

$$1 + \lambda = \prod_{j=1}^{n} (1 + \lambda g^j). \quad (6)$$

Sugeno-type measures are widely applied in the context of fuzzy reasoning and aggregation, particularly when modeling interactions among dependent criteria (Beliakov *et al.*, 2007).

Based on such fuzzy measures, the *Choquet integral* provides a powerful aggregation mechanism. We recall its discrete formulation below.

Definition 2. Let g be a fuzzy measure and let $h : X \to [0, 1]$ be a function such that $h(z_1) \geq h(z_2) \geq \cdots \geq h(z_n)$. The *discrete Choquet integral* of h with respect to g is defined as

$$C_g(h) = \sum_{i=1}^{n} (h(z_i) - h(z_{i+1})) \cdot g(A_i), \quad (7)$$

where $A_i = \{z_1, \ldots, z_i\}$ and $h(z_{n+1}) = 0$. The ordering of h ensures antitonicity.

In our federated learning framework, we use the Choquet integral to aggregate parameters learned by local models. To reflect the uncertainty inherent in each model's parameters, we introduce a generalized entropy-based weighting mechanism built upon *interval-valued fuzzy sets*.

The following definition introduces the concept of *interval entropy*, as proposed by Pękala *et al.* (2025), which enhances earlier definitions such as that by Zeng and Guo (2008) through explicitly incorporating the width of intervals to better reflect the degree of uncertainty.

Definition 3. Let N be a strong (involutive) interval negation with an equilibrium point $e \in L^I$ such that $N(e) = e$ and e is closest to $[0.5, 0.5]$. A function E_{INT} : IVFS$(X) \to L^I$ is called an *interval entropy* with respect to N if it satisfies the following axioms for all $S, T \in$ IVFS(X):

(e1) $E_{INT}(S) = 0_{L^I}$ if and only if S is a crisp (non-fuzzy) set;

(e2) $E_{INT}(\mathbf{E}) = [1 - w(e), 1]$, where $w(e)$ denotes the width of the equilibrium point;

(e3) $E_{INT}(S) \leq E_{INT}(T)$ if $S(x) \leq T(x) \leq e$ or $S(x) \geq T(x) \geq e$ for all $x \in X$;

(e4) $E_{INT}(S) = E_{INT}(S_N)$, where $S_N(x) = N(S(x))$ and $\mathbf{E}(x) = e$.

This interval entropy provides a consistent and interpretable measure of uncertainty, which we integrate into the aggregation process in federated learning. Specifically, it allows the Choquet integral to account not only for model performance (e.g., via the AUC) but also for the associated imprecision in learned parameters, thereby enabling a more robust fusion of knowledge across clients.

3.2.1. Entropy construction via a similarity measure. In the following proposition, we present how an appropriate similarity function can be employed to construct an interval-valued entropy measure. This approach aligns with the axiomatic framework introduced earlier and provides a practical computation scheme based on similarity between interval-valued fuzzy sets (Pękala *et al.*, 2021).

Proposition 1. *Let the* SIM *be an interval similarity measure satisfying the conditions introduced by Pękała et al. (2021), where A_1 denotes an idempotent aggregation function. Then, the function E_{INT} : $\mathrm{IVFS}(X) \rightarrow L^I$ defined as*

$$E_{INT}(S) = \mathrm{SIM}(S, S_N) \tag{8}$$

is an interval-valued entropy measure with respect to the involutive interval negation N having an equilibrium point e.

3.2.2. Application in federated model aggregation. We now explain how the interval entropy and the Choquet integral are integrated to perform parameter aggregation across local models in a federated learning setting.

In the proposed methodology, the density functions and the parameter λ in the Choquet framework are derived based on local model performance and their associated uncertainty. Let r denote the number of clients (local models), and Q_i be the quality measure of the i-th local model. In our case, $Q_i = \mathrm{AUC}_i$, i.e., the area under the ROC curve, as it effectively captures a classifier's discriminative ability based on sensitivity and specificity.

We then define the fuzzy density $g(x_i)$ and determine λ as follows:

1. If $\sum_{i=1}^{r} Q_i \leq 1$, then

$$g(x_i) = \frac{|Q_i - E_{INT_i}|}{\sum_{i=1}^{r} Q_i}, \qquad \lambda = 0.$$

2. If $\sum_{i=1}^{r} Q_i > 1$, then

$$g(x_i) = |Q_i - E_{INT_i}|$$

and λ is calculated from Eqn. (6), while the values of $g(\{x_i\})$ for all subsets are recursively computed using Eqn. (5).

The collective entropy E_{INT_i} for the i-th client is calculated by aggregating entropy values across all n objects in that client's dataset:

$$E_{INT_i} = Rep_{0.5}\left(\mathcal{A}_{l=1,\ldots,n}(E_{INT_{i_l}})\right),$$

where $E_{INT_{i_l}}$ is the entropy of the l-th object in client i, and $Rep_{0.5}(x)$ denotes the midpoint representation of an interval:

$$Rep_{0.5}(x) = \frac{\underline{x} + \overline{x}}{2}.$$

3.2.3. Choquet-based model parameter aggregation. The final aggregation of local model parameters is performed using the Choquet integral. Let $x = (\rho_i^1, \ldots, \rho_i^r)$ represent the i-th parameter across all r clients, where all parameters are first normalized.

We compute the aggregated parameter as

$$\rho_i^{\text{agg}} = C_g(h),$$

where $h(x_i) \in \{\rho_i^1, \ldots, \rho_i^r\}$ and $C_g(\cdot)$ is calculated according to Eqn. (7). This results in a global model that fuses local knowledge with consideration for both quality and uncertainty.

4. Experimental results and discussion

Here, we concentrate on two directions of study, i.e., firstly, we evaluate and compare the proposed *interval-weights learning method* with classical approaches to learn weights; secondly, we examine the performance of federated aggregation using the *Choquet integral with entropy-based fusion*.

The interval-weight learning technique introduced by Szkoła *et al.* (2024) serves as the core of our local model training process in both paths.

The choice of the Choquet integral as the aggregation method in the second research path is dictated by the results obtained in our previous articles/research, where we demonstrated that the Choquet integral is a more efficient aggregation method than weighted averages based on various weighting techniques, in particular, generalizing the FedAVG method based on the number of objects per client to other sources of information about the models used for weights, such as model efficiency. Hence, demonstrating improvement over the

Choquet-based aggregation also implies progress beyond classical averaging techniques, such as arithmetic or weighted means.

Therefore, the experimental framework presented in this section builds upon earlier studies (Wilbik *et al.*, 2023; 2022; Pękala *et al.*, 2024), which investigated the effectiveness of various weighted aggregation schemes in federated learning settings, and explains the motivation for selecting the Choquet integral, which stems from its demonstrated effectiveness as a fusion operator in heterogeneous environments (Pękala *et al.*, 2024; 2025).

Evaluated models and setup. Our experiments investigate the following six model configurations:

- *Model 1*: centralized model using the entire dataset and classical learning;
- *Model 2*: centralized model with the interval-weights learning method;
- *Model 3*: local models trained independently with classical learning;
- *Model 4*: local models trained with interval-weights learning;
- *Model 5*: federated learning with classical learning and aggregation via the Choquet integral with entropy;
- *Model 6*: federated learning using interval-weights learning and the Choquet integral with entropy-based fusion.

To simulate a realistic FL scenario, the dataset was randomly partitioned into three disjoint groups, each representing a distinct client. The use of three clients reflects a realistic multi-institutional configuration while preserving sufficient data per node for statistical reliability. With larger datasets, our framework scales naturally to a higher number of clients, as no constraint is imposed on client count. For each client, the dataset was split into a training set (90%) and a test set (10%).

Across all models, we employed logistic regression as the base classifier, configured with the following hyperparameters: $\epsilon_i = 0.001$ (regularization term), $\alpha = 0.01$ (learning rate), and $\gamma = 0.5$ (interval representation parameter).

Parameter settings and data distribution. To study the sensitivity of our approach, we varied the number of subintervals used in the interval encoding of model parameters, considering values of $z \in \{2, \ldots, 7\}$. This allowed us to investigate the granularity-accuracy trade-off inherent in interval-based modeling.

Table 1. Model 1: centralized model with classical learning.

Dataset	ACC	SENS	SPEC	PREC
Complete	0.965	0.972	0.935	0.965

Table 2. Model 2: centralized model with interval-weighted learning.

Dataset	ACC	SENS	SPEC	PREC
Complete	**0.983**	0.972	**0.950**	**0.970**

For data distribution among clients, we simulated a *non-IID* scenario, where class imbalance was introduced intentionally. The distribution was as follows:

- *Client 1*: balanced dataset, i.e., 50% malignant, 50% benign;
- *Client 2*: unbalanced, i.e., 70% malignant, 30% benign;
- *Client 3*: unbalanced, i.e., 70% benign, 30% malignant.

This heterogeneity mirrors real-world conditions where participating institutions hold datasets with varying class distributions, posing major challenges for collaborative learning.

4.1. Results and discussion across scenarios.

Evaluation metrics. To assess the performance of our models, we adopt standard classification metrics, defined as follows:

- accuracy (ACC): $ACC = \frac{TP+TN}{TP+TN+FP+FN}$,
- specificity (SPEC): $SPEC = \frac{TN}{TN+FP}$,
- precision (PREC): $PREC = \frac{TP}{TP+FP}$,
- sensitivity (SENS): $SENS = \frac{TP}{TP+FN}$.

Here, TP denotes the number of true positives (malignant cases correctly classified), TN the number of true negatives (benign cases correctly classified), FP the number of false positives (benign cases misclassified as malignant), and FN the number of false negatives (malignant cases misclassified as benign).

We also incorporate the *area under the curve* metric as part of the model aggregation process via the Choquet integral. The AUC is widely used to evaluate classifier discriminative ability, as it reflects the area under the ROC curve—a plot that connects sensitivity and specificity across varying thresholds.

We begin by analyzing the performance of centralized models. Table 1 presents results for Model 1,

Table 3. Model 3: local models with classical learning (NON-IID).

Client	ACC	SENS	SPEC	PREC
1	0.94	0.96	0.89	0.94
2	0.81	0.86	0.89	0.89
3	0.70	0.89	0.78	0.88

Table 4. Model 4: local models with interval-weighted learning (NON-IID).

Client	ACC	SENS	SPEC	PREC
1	0.94	0.96	0.89	0.94
2	**0.91**	0.86	**0.99**	**0.99**
3	**0.78**	**0.97**	0.65	0.81

which uses classical logistic regression with real-valued weights. As a point of reference, Model 2, shown in Table 2, employs the proposed interval-weighted learning scheme. Notably, the interval-based method yields improvements across all evaluation metrics, especially in precision and specificity.

These results clearly indicate that incorporating interval-valued weights improves model stability by better representing uncertainty inherent in the training data, even when access to the full dataset is available.

4.1.1. Local models: Classical vs interval-weighted learning. We now turn to the federated scenario, in which local clients operate on disjoint, non-IID subsets of the data. Table 3 shows the performance of Model 3, where clients train models independently using the classical method. Due to class imbalance, Clients 2 and 3 receive predominantly malignant and benign cases, respectively, and model quality degrades substantially.

The results demonstrate that traditional local training fails to generalize well under data skew, which is common in real-world FL deployments.

Next, Table 4 shows results for Model 4, where local training incorporates interval-weighted learning. Even without federation, the method enhances performance for Clients 2 and 3, most notably improving specificity and precision—key metrics in medical decision support.

These improvements highlight the method's ability to compensate for partial or skewed data through richer internal representation of feature uncertainty.

4.1.2. Federated models: Choquet aggregation with entropy. Table 5 summarizes the results for Model 5, where classical learning is used locally but the fusion is carried out using the Choquet integral with entropy-based weighting. Despite no modification in the local model structure, the performance of all clients improves due to the entropy-aware aggregation, which downweights unreliable models and emphasizes informative ones.

Table 5. Model 5: FL with Choquet aggregation and classical learning.

Client	ACC	SENS	SPEC	PREC
1	0.96	0.94	0.99	0.99
2	0.92	0.89	0.95	0.97
3	0.92	0.97	0.86	0.92

Table 6. Model 6: FL with Choquet aggregation and interval-weighted learning.

Client	ACC	SENS	SPEC	PREC
1	0.96	0.94	0.99	0.99
2	**0.97**	**0.95**	**0.99**	**0.99**
3	0.92	0.97	**0.87**	0.92

Finally, in Table 6, we present the results of Model 6, the most advanced setup combining interval-weighted learning with Choquet entropy aggregation. This dual-layered uncertainty handling achieves the highest overall performance across clients.

4.1.3. Discussion. The experimental results presented above lead to several important insights regarding the role of uncertainty modeling in federated learning. In the centralized scenario, we observe that incorporating interval-valued weights into the logistic regression model leads to consistent improvements across evaluation metrics, particularly in terms of precision and specificity. This indicates that interval representations are better equipped to capture ambiguity and borderline instances—a feature especially relevant in clinical or diagnostic data.

In non-IID federated settings, local models trained with classical learning exhibit notable performance degradation, particularly for clients with skewed class distributions. While standard aggregation partially mitigates these effects, it cannot fully recover performance unless the aggregation process explicitly incorporates the uncertainty of individual models.

The proposed entropy-aware Choquet aggregation mechanism addresses this limitation by weighting each client's contribution based on both model quality and confidence. This enables the global model to down-weight information from unreliable or unbalanced sources while enhancing the influence of clients providing more stable and informative updates. The result is a more robust and balanced global model, less prone to overfitting toward dominant local distributions.

The most substantial gains are achieved through the combined application of interval-weighted learning and entropy-based fusion, as demonstrated by Model 6. This approach consistently outperforms all other models, including the centralized baseline, particularly for clients exposed to highly unbalanced data. These findings confirm that simultaneous modeling of uncertainty at both the local and global levels can effectively address the dual challenges of heterogeneity and distributional bias inherent to federated learning.

The relevance of this approach is particularly evident in medical applications, where high specificity and precision are critical for avoiding false positives. The interval-based modeling captures data uncertainty at the feature level, while the entropy-driven fusion ensures that less reliable sources do not dominate the learning process. Together, these mechanisms yield improved diagnostic reliability across clients, supporting the method's potential utility in distributed healthcare AI.

It should be noted, however, that the proposed methodology incurs additional computational overhead due to interval operations and entropy calculations. Nevertheless, the observed improvements in performance strongly justify further development of this line of research. Future work may explore adaptive determination of subinterval structures, entropy regularization strategies, and personalized optimization schedules for individual clients to enhance scalability and efficiency.

In terms of computational complexity, the total cost of the proposed approach can be approximated by

$$O(m \cdot E \cdot n \cdot P \cdot C + m \cdot P \cdot T), \tag{9}$$

where m is the number of clients, E denotes local epochs, n stands for the dataset size per client, P is the number of model parameters, C denotes the cost of training one instance, and T stands for the cost of aggregating parameter vectors. Although interval and entropy computations introduce moderate overhead, comparative tests with FedAvg indicate less than 20% additional cost, compensated by substantial performance gains.

5. Summary and future work

This paper presented a novel methodology for federated learning in uncertain and heterogeneous environments. By integrating interval-valued fuzzy sets into the local training process and applying the Choquet integral with entropy for aggregation, we developed an interpretable, uncertainty-aware learning pipeline.

The proposed framework enhances logistic regression with interval-weighted parameters, effectively capturing imprecision at both local and global levels. Experimental results demonstrate superior performance, especially under class imbalance, compared to traditional learning and aggregation methods.

Future work will focus on extending the interval aggregation scheme by incorporating higher-order entropy and adapting the approach to practical applications, such as, e.g., fall detection in smart environments. Additionally, we aim at developing new strategies for automated weighting in Choquet-based fusion, improving adaptability and scalability in real-world federated systems.

The main limitation of this study lies in the use of a single medical dataset, which restricts empirical generalization. Future work will therefore include evaluation on diverse datasets and domains (e.g., sensor-based monitoring), as well as optimization of interval operations for large-scale federated deployments.

References

Beliakov, G., Pradera, A. and Calvo, T. (2007). *Aggregation Functions: A Guide for Practitioners*, Studies in Fuzziness and Soft Computing, Vol. 221, Springer, Berlin/Heidelberg, pp. 1–37.

Bustince, H., Barrenechea, E., Pagola, M., Fernandez, J., Xu, Z., Bedregal, B., Montero, J., Hagras, H., Herrera, F. and De Baets, B. (2016). A historical account of types of fuzzy sets and their relationships, *IEEE Transactions on Fuzzy Systems* **24**(1): 179–194.

Choquet, G. (1953). Theory of capacities, *Annales de I'Institut Fourier* **5**: 131–295.

Deist, T.M., Jochems, A., van Soest, J., Nalbantov, G., Oberije, C., Walsh, S., Eble, M., Bulens, P., Coucke, P., Dries, W., Dekker, A. and Lambin, P. (2017). Infrastructure and distributed learning methodology for privacy-preserving multi-centric rapid learning health care: EuroCAT, *Clinical and Translational Radiation Oncology* **4**: 24–31.

Gorzałczany, M.B. (1987). A method of inference in approximate reasoning based on interval-valued fuzzy sets, *Fuzzy Sets and Systems* **21**(1): 1–17.

Hard, A., Kiddon, C.M., Ramage, D., Beaufays, F., Eichner, H., Rao, K., Mathews, R. and Augenstein, S. (2018). Federated learning for mobile keyboard prediction, `https://arxiv.org/abs/1811.03604`.

Kairouz, P., McMahan, B., Avent, B., Bellet, A. and Bennis, M. (2021). Advances and open problems in federated learning, *Foundations and Trends® in Machine Learning* **14**(1): 1–210.

Konecný, J., McMahan, H.B., Ramage, D. and Richtárik, P. (2016). Federated optimization: Distributed machine learning for on-device intelligence, *ArXiv* 1610.02527.

Konečný, J., McMahan, H.B., Yu, F.X., Richtárik, P., Suresh, A.T. and Bacon, D. (2017). Federated learning: Strategies for improving communication efficiency, *arXiv* 1610.05492.

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.

McMahan, H.B., Moore, E., Ramage, D., Hampson, S. and y Arcas, B.A. (2017). Communication-efficient learning of deep networks from decentralized data, *AISTATS 2017, Ft. Lauderdale, USA,* pp. 1273–1282.

Piegat, A. and Landowski, M. (2013). Multidimensional approach to interval uncertainty calculations, *in* K. Atanassov *et al.* (Eds), *New Trends in Fuzzy Sets, Intuitionistic: Fuzzy Sets, Generalized Nets and Related Topics. Volume II: Applications*, IBS PAN–SRI PAS, Warsaw, pp. 137–151.

Pękala, B., Grzegorzewski, P., Szkoła, J. and Kosior, D. (2025). PRISM-FL—A privacy-aware decision support system for enhanced medical diagnostics based on federated learning, *IEEE Access* **13**: 1–13.

Pękala, B., Wilbik, A., Szkoła, J., Dyczkowski, K. and Żywica, P. (2024). Federated learning with the Choquet integral as aggregation method, *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Yokohama, Japan*, pp. 1–8.

Sambuc, R. (1975). *Fonctions ϕ-floues: Application á l'aide au diagnostic en pathologie thyroidienne*, PhD thesis, University of Marseille, Marseille.

Szkoła, J., Pękala, B. and Dyczkowski, K. (2024). Refining uncertainty management in machine learning: An interval-valued fuzzy set approach to logistic regression, *Information Processing and Management of Uncertainty in Knowledge-Based Systems: Theory and Methods, IPMU 2024, Lisbon, Portugal*, pp. 1–14.

Türksen, I.B. (1986). Interval valued fuzzy sets based on normal forms, *Fuzzy Sets and Systems* **20**(2): 191–210.

UCI (2017). *UCI Machine Learning Repository*, University of California, Irvine, `http://archive.ics.uci.edu/ml`.

Wilbik, A. and Grefen, P. (2021). Towards a federated fuzzy learning system, *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Luxembourg*, pp. 1–6.

Wilbik, A., Pękala, B., Dyczkowski, K. and Szkoła, J. (2022). A comparison of client weighting schemes in federated learning, *International Workshop on Intuitionistic Fuzzy Sets and Generalized Nets, Warsaw, Poland*, pp. 116–128.

Wilbik, A., Pękala, B., Szkoła, J. and Dyczkowski, K. (2023). The Sugeno integral used for federated learning with uncertainty for unbalanced data, *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE'2023, Songdo Incheon, Korea*, pp. 1–6.

Yan, H., Hu, L., Xiang, X., Liu, Z. and Yuan, X. (2021). Privacy-preserving collaborative learning for mitigating indirect information leakage, *Information Sciences* **548**: 423–437.

Yang, Q., Liu, Y., Chen, T. and Tong, Y. (2019). Federated machine learning: Concept and applications, *ACM Transactions on Intelligent Systems and Technology* **10**(2): 1–19.

Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning—I, *Information Sciences* **8**(3): 199–249.

Zeng, W. and Guo, P. (2008). Normalized distance, similarity measure, inclusion measure and entropy of interval-valued fuzzy sets and their relationship, *Information Sciences* **178**(5): 1334–1342.

Zheng, W., Yan, L., Gou, C. and Wang, F.-Y. (2020). Federated meta-learning for fraudulent credit card detection, *Proceedings of the 29th International Joint Conference on Artificial Intelligence (IJCAI-20), Yokohama, Japan,* pp. 4654–4660.

Jarosław Szkoła works at the University of Rzeszów, Institute of Computer Science. His research interests focus on modeling and developing neural network architectures, analysis of incomplete data using machine learning algorithms, and using federated learning. He has authored many didactic tools in the form of decision games. He also has extensive experience in designing electronic devices, simple mobile robots, sensory data acquisition, and processing systems. ORCID: 0000-0002-6043-3313.

Barbara Pękala is the dean of the Faculty of Exact and Technical Sciences, University of Rzeszów, Poland. Her research interests concentrate on the exploration and application of extensions of fuzzy set theory, e.g., aggregation functions, information, and comparison measures in computational intelligence methods for decision-making, such as approximate reasoning, image processing, data mining, and federated learning systems. She is a member of the Fuzzy Systems Technical Committee (FSTC) of the Computational Intelligence Society (CIS), the president of the Polish Society of Fuzzy Sets (POLFUZZ), and a vice-president of EUSFLAT. ORCID: 0000-0002-5501-5467.

Krzysztof Dyczkowski is the dean of the Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Poznań, Poland. His main research interests are computational intelligence methods in decision-making, computer vision and natural language modeling. He is a recipient of the AMU Foundation Award Excellence in Science—Usefulness in Practice (2019), the Microsoft Research Azure Award (2015) and the IBM Faculty Award (2010). He is a member of the IEEE Computational Intelligence Society and EUSFLAT, and a vice-president of the Polish Society of Fuzzy Sets (POLFUZZ). ORCID: 0000-0002-2897-3176.

Received: 31 July 2025
Accepted: 3 November 2025