JCSE Journal of Computing Science and Engineering

Regular Paper
Journal of Computing Science and Engineering,
Vol. 19, No. 1, March 2025, pp. 16-23

A Federated Learning Method for Multi-Positive and Unlabeled Data Using an Ensemble of One-Positive and Unlabeled Learning Models

Cheong Hee Park*
Department of Computer Science and Engineering, Chungnam National University, Daejeon, Korea
cheonghee@cnu.ac.kr

Abstract
Centralized learning, which requires data to be collected in a central location, is ill-suited to meet the requirements of stricter regulations on data privacy and security. Federated learning provides solutions to this problem, enabling a model to be trained across a group of clients while maintaining the privacy of each client's data. When a client has small-sized labeled data from a part of the classes with large-sized unlabeled data, referred to as a multi-positive and unlabeled learning problem, leveraging unlabeled data can improve the performance of the global model aggregated from the local models of the clients. In this paper, we propose a federated learning method for multi-positive and unlabeled data using an ensemble of one-positive and unlabeled learning models. We construct an ensemble of positive and unlabeled learning models for each positive class in clients, which is trained as a binary classifier on one positive class with small-sized labeled data and one negative class consisting of all remaining data. The ensemble collected from all clients is used to extend the labeled data of each class in clients. The experimental results using image and text data show that the proposed method improved the performance for small-sized multi-positive and unlabeled federated learning.

Category: Artificial Intelligence

Keywords: Extension of labeled data; Federated learning; Multi-positive and unlabeled learning; Semi-supervised federated learning; Positive and unlabeled learning

I. INTRODUCTION

Traditionally, a machine learning model is expected to learn from data stored in a central location, but with stricter regulations on data privacy and safety, the paradigm is shifting from centralized learning to federated learning. Instead of storing all data in a centralized location, federated learning involves training a machine learning model collaboratively with participants with the collected data kept locally, preserving data privacy [1, 2]. Federated learning has been a hot research topic recently with applications in various areas such as medical, financial, and industrial domains [3-5].

In federated learning, a server keeps a global model and updates it iteratively by sending parameters of the global model to participating clients, receiving the parameters of the models updated locally in clients, and aggregating them [6]. However, when the data distribution in clients is non-independent and identically distributed (non-IID), local models on clients can diverge and their aggregated model may require additional rounds of communication to reach the expected performance level

Open Access http://dx.doi.org/10.5626/JCSE.2025.19.1.16 http://jcse.kiise.org

Received 16 October 2024; **Accepted** 30 May 2025
*Corresponding Author

 pISSN: 1976-4677 eISSN: 2093-8020

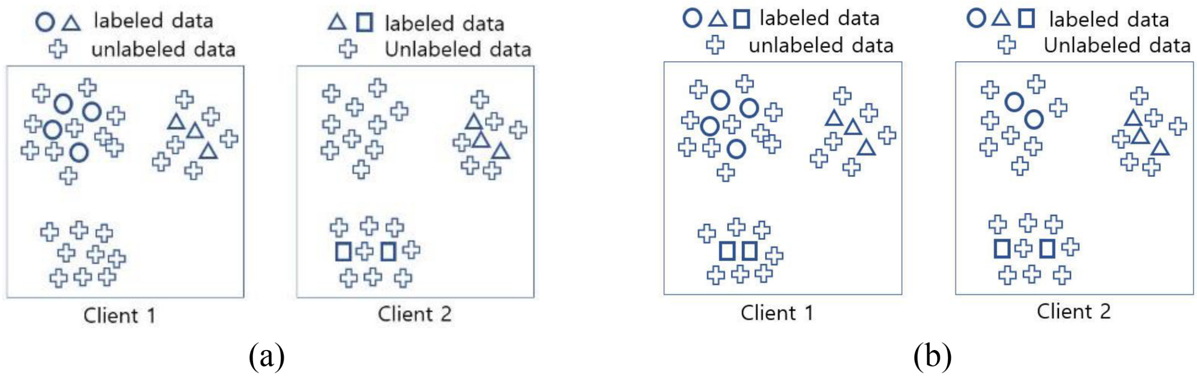

Fig. 1. Illustration of (a) MPUL and (b) semi-supervised learning.

[2, 7, 8].

Moreover, clients can have small-sized labeled data belonging to a part of classes and large-sized unlabeled data. This problem is called multi-positive and unlabeled learning (MPUL), while positive and unlabeled learning generally refers to a learning problem on labeled data from one positive class and unlabeled data from the positive and the negative classes. MPUL is also different from semi-supervised learning which generally assumes a situation where small labeled data and unlabeled data exist for each class. Fig. 1 illustrates the examples of MPUL and semi-supervised learning. In Fig. 1(a), which describes MPUL, client 1 has labeled data of class O, Δ, but client 2 has labeled data of class □, Δ. In Fig. 1(b), for semi-supervised learning, two clients have labeled data from all three classes O, Δ, □.

Recently, Lin et al. [9] proposed FedPU, a federated learning method for MPUL, where information for negative classes without labeled data in one client is derived from labeled data of those classes in other clients. They showed that FedPU can achieve higher performance compared to other semi-supervised federated learning methods. However, when the amount of labeled data of positive classes in clients is small, it cannot give enough information for the negative classes of other clients. In this paper, we propose a method for federated learning with small-sized multi-positive data and large unlabeled data. We construct an ensemble of positive and unlabeled (PU) learning models which are trained for each positive class in clients. A simple PU learning model is trained as a binary classifier on one positive class with labeled data and one negative class consisting of all remaining labeled and unlabeled data. The ensemble is used to extend the labeled data of each class in clients. Then any semi-supervised federated learning models can be applied to the extended labeled data and the remaining unlabeled data. Fig. 2 illustrates the proposed method. The server collects two PU models, f_o^1 and f_Δ^1, which are trained for positive class O and Δ in client 1, respectively and two PU models, f_Δ^2 and $f_\square^2$, which are trained for positive class Δ and □ in client 2, and the ensemble $\{f_o^1, f_\Delta^1, f_\Delta^2, f_\square^2\}$ is sent to all clients to expand reliable labeled data for all the classes in clients. Experimental results using image and test data show that the proposed method gives the improved performance for small-sized multi-positive and unlabeled federated learning.

The rest of this paper is organized as follows. Section

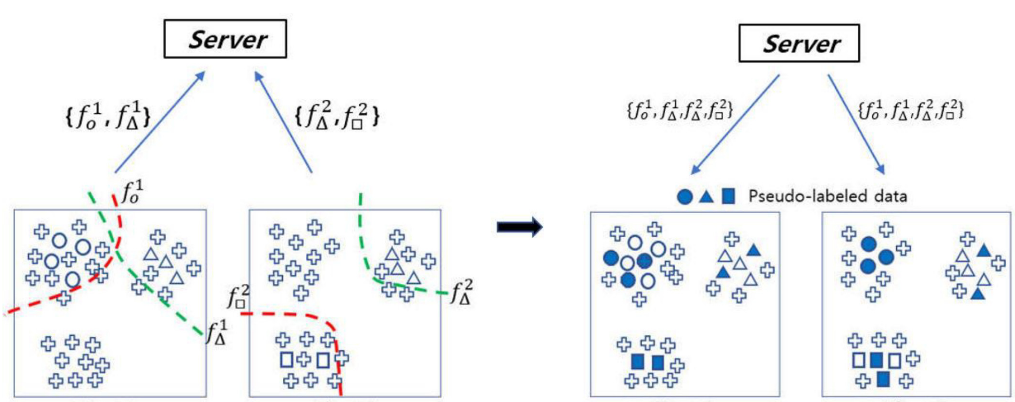

Fig. 2. An illustration of the proposed method.

II provides an overview of relevant background concepts and related research. Section III presents our proposed method. Section IV introduces experimental settings and results followed by a conclusion in Section V.

II. RELATED WORK

A. Positive and Unlabeled Learning

PU learning aims to learn a classifier when labeled data from a positive class and unlabeled data from both a positive class and an unknown negative class are given [10]. Many PU learning methods adopt a two-step approach to identify reliable negative data from unlabeled data and learn a binary classifier. In [11], “spy” documents were randomly selected from the positive class and moved to the unlabeled dataset. Initializing all unlabeled data as negative data, reliable negative documents in unlabeled data were determined by using a naive Baysian classifier and expectation-maximization (EM) algorithm where the probabilistic labels of the spies are used to determine the most likely negative documents. Finally, the EM algorithm is again applied for classification with positive, negative, and unlabeled data. In [12], graph-based label propagation was applied after extracting reliable negative documents based on the average distance from all positive documents. Various risk minimization formulations were also applied for PU learning, such as convex loss minimization using different loss functions [13], a non-negative risk estimator [14], and loss decomposition and estimation of negative data centroid regarding unlabeled data as noisy negative data [15].

While PU learning is based on a binary classification, MPU assumes that labeled data from multiple positive classes and unlabeled data from either the positive classes or an unknown negative class are given [16, 17]. MPU can arise in various real problems. In fraud transaction detection, where fraud transaction data can be considered as positive data, there are multiple fraud types [17], and in document classification, labeled positive data can be composed of several categories of documents. In [16], using different loss functions for labeled data and unlabeled data, the original data space is mapped onto an embedding space where the codewords corresponding to each class are fixed by a maximum margin. The parameter matrix of a linear discriminant function and the label estimation of unlabeled data samples are alternatively optimized. In [17], to address the overfitting problem caused by the unbounded risk of the method of [16], an alternative risk estimator with the modification of the hinge loss function has been proposed. In [18], the original data space is mapped onto a low dimensional space by a linear discriminant function and reliable negative data samples are selected by using an ensemble of KNN-based outlier detection models in low dimensional data space. Parameter updates of a linear discriminant function and the selection of reliable negative samples by outlier detection in low dimensional space are alternatively performed.

B. Federated Positive and Unlabeled Learning

Recently, in [9], a federated learning method with multi-positive and unlabeled data, FedPU, was proposed, where not only positive data in a client consists of multiple classes, but also negative data may come from multiple classes which are unknown to the client. In FedPU, from the fact that the weights in the global model are derived from the combination of local models updated in each client, the expected risk for negative classes in one client is minimized by leveraging the labeled data of those classes in other clients. However, when the amount of positive data is small, it will not give enough information for the negative classes of other clients and the performance can be severely degraded. In the next section, we present a method for federated learning on small-sized multi-positive data and large unlabeled data, which performs the extension of labeled data by the ensemble of simple PU learning models.

III. THE PROPOSED METHOD

When the set of positive classes to which the labeled data of client k (k=1,2,...,K) belongs is represented by C_{P_k} and the set of negative classes that do not contain labeled data is represented by C_{N_k}, we assume that unlabeled data come from the entire class label set $C = C_{P_k} \cup C_{N_k}$ Extraction of reliable pseudo-labeled data in the proposed method is performed in a two-step approach. First, in a client, a positive and unlabeled learning model for each positive class with labeled data is constructed by binary classification on labeled data of the positive class versus all the remaining data. These models constructed in all clients are sent to the server and then they are assembled as an ensemble. Secondly, the ensemble of the models is sent back to all the clients where the extension of labeled data for all classes is performed. Then, any semi-supervised federated learning model can be applied with the extended labeled data and remaining unlabeled data in clients. In the following subsections, the two steps in the proposed method are explained in detail.

A. Construction of a PU Learning Model for Each Positive Class with Labeled Data in a Client

Many PU learning methods apply various strategies to extract reliable negative data among unlabeled data. One simple approach is to assume all unlabeled data is data of a negative class and apply a binary classifier to select

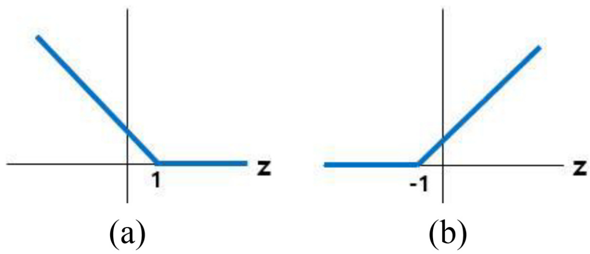

Fig. 3. Illustrations of the hinge loss $h(z) = \max\{0, 1 - z\}$: (a) $h(z)$ and (b) $h(-z)$.

data samples predicted to a negative class with high confidence. We adopt this approach to construct a classifier for each positive class with labeled data.

Let $F(x) = [f_1(x), f_2(x)]$ be the output in a binary classification model F where one-hot encoded target values for positive data are set as [1,0] and target values for negative data are [0,1]. The objective function by the hinge loss $h(z) = max\{0, 1-z\}$ shown in Fig. 3 is defined as,

$$\text{minimize } \frac{1}{|D_1|}\sum_{x \in D_1}[h(f_1(x)) + h(-f_2(x))] + \alpha\frac{1}{|D_2|}\sum_{x \in D_2}[h(-f_1(x)) + h(f_2(x))] \quad (1)$$

The first term in Eq. (1) is the loss on positive data and the second term is for the loss on negative data. α is a parameter which controls the weight of the loss of negative data over the loss of positive data. In the experiments of Section IV, we used a neural network as a classifier for PU learning, adjusting the value of α adaptively during learning epochs.

For each class i in the positive classes C_{P_k} of client k, considering the labeled data of class i as positive data and all the remaining data as negative data, the model F_i^k is learned using the objective function in Eq. (1). Since only a small part of data in the positive class i is labeled, all unlabeled data of class i is set as negative data with data from other classes. Hence, the loss in the second term could be very big in the initial learning rounds and the decision boundary may be determined defensively for the positive class i. We set the parameter α as 1 initially and change it adaptively so as to secure the desirable size of data samples predicted to positive class i. After each epoch of training, class labels for training data are predicted. If the number of data samples predicted to positive class i is less than the predefined *num_limit*, α is decreased as $\alpha * 0.9$. If the number of data samples predicted to positive class i is more than the total number of local data samples divided by |C|, α is increased as min{1, α/0.9. This building process is summarized in *BuildPU( )* of Table 1. The models constructed for each class in C_{P_k} in the client k, $\{F_i^k | i \in C_{P_k}\}$, are sent to the server.

B. Extension of Labeled Data for Each Class using the Ensemble of PU Models

The server builds an ensemble $E = \{F_i^k | 1 \le k \le K,\ i \in C_{P_k}\}$ of PU models uploaded from clients and sends the ensemble back to all the clients. Clients extend labeled data based on the prediction by the ensemble E. Let $\tilde{F}_i^k(x) = [\tilde{f}_{i1}^k(x), \tilde{f}_{i2}^k(x)]$ be the softmax function value on output $F_i^k(x) = [f_{i1}^k(x), f_{i2}^k(x)]$ of a PU learning model for positive class i. $\tilde{f}_{i1}^k(x)$ gives the confidence for the prediction of x to the class i by F_i^k. We define the confidence for the prediction of x to the class i as,

$$u_i(x) = \max\{\tilde{f}_{i1}^k(x) | 1 \le k \le K,\ i \in C_{P_k}\} \quad (2)$$

and the data sample x is predicted to class j which is $j = \text{argmax}_i\{u_i(x)\}$ with confidence $u_j(x)$.

We add pseudo-labeled data to each class by selecting unlabeled data samples which have high confidence for the prediction to that class. The number of data samples added to each class can be limited by the parameter *num_add*. In the experiments of Section IV, we set the value of *num_add* as *num_limit/2*. Now, based on the extended labeled data, any semi-supervised federated learning method can be applied. The proposed method is summarized in Table 1.

IV. EXPERIMENTAL EVALUATION

A. Experimental Results using Text Datasets

We evaluated the performance of our method using three text datasets: Reuters, Sports, and Classic. Reuters-21578 was downloaded from UCI Machine Learning Repository and the documents belonging to 135 TOPICS categories were used. After preprocessing by stopwords removal, stemming, term frequency-inverse document frequency (TF-IDF) transformation, and unit norm, we had 9,805 documents consisting of 15,484 words. Three classes were composed of the two largest categories of 1 and 36 and the collection of all the remaining documents. Sports and Classic datasets were downloaded from [19]. They were constructed by removing classes with less than 1,000 texts and terms with frequencies less than or equal to 1. The detailed description for datasets is shown in Table 2.

Each dataset was split to training and test data with an 8:2 ratio. Since the number of classes in the text data is small, we set the number of clients to be the same as the number of classes and uniformly divided data of each class in training data to clients. Each class was designated as a positive class in one of the clients and t percent of the

Table 1. The proposed method for multi-positive and unlabeled federated learning

Input:
K: the number of clients
C_{P_k}: the set of positive classes to which the labeled data of client k belongs
C_{N_k}: the set of negative classes of client k that do not contain labeled data
$C = C_{P_k} \cup C_{N_k}$: the entire class label set from which unlabeled data of client k come
num_limit, num_add: parameters for *buildPU( )* and *Extend( )*

1. Each client $k(1 \le k \le K)$ builds PU learning model F_i^k for each class i in C_{P_k} by *buildPU( )* and sends $\{F_i^k | i \in C_{P_k}\}$ to the server
2. Server assembles $E = \{F_i^k | 1 \le k \le K,\ i \in C_{P_k}\}$ and sends E to all the clients.
3. Each client $k(1 \le k \le K)$ extends the set of labeled data in its local data based on the prediction by the ensemble E using *Extend( )*.

buildPU( ) in client k

```
1. for each class i ∈ C_{P_k}
2.    let D_1 be the set of labeled data of class i and D_2 be the set of all the remaining data
3.    Initialize α = 1 and the parameters of the model F_i^k for the binary classification on the dataset D_1 of class 1 and
      the dataset D_2 of class 2
4.    for local epochs
5.       update the parameters of the model F_i^k using the objective function in Eq. (1)
6.       Predict the class labels on data of D_1 ∪ D_2 by F_i^k and let M be the number of data samples predicted to class 1
7.       α ← α * 0.9 if M < num_limit, and α ← min{1, α/0.9} if M > |D_1 ∪ D_2| / |C|
8.    end for
9. end for
```

Extend( ) in client k

```
1. for unlabeled data sample x of client k
2.    for each class i ∈ C
3.       compute the confidence u_i(x) for the prediction of x to the class i by Eq. (2)
4.    end for
5.    Predict x to class j which is j = argmax_i{u_i(x)} with confidence u_j(x)
6. end for
7. for each class i ∈ C
8.    Extend the labeled dataset of class i with unlabeled data samples predicted to class i with high prediction
      confidence whose added size is limited by num_add
9. end for
```

Table 2. Data description

Data	Dim	Samples	Classes
Reuters	15,484	9,805	3
Sports	18,324	7,168	3
Classic	12,009	7,094	4

data in the positive class was labeled. We conducted different settings of $t = 1\%$ or 2% and the test was repeated 10 times with random data splitting. We evaluated the performance of our proposed method by comparing the accuracy of the three methods: the supervised learning model (SL), the MPU federated learning model in [9] (FedPU), and the MPU federated learning model after the

Table 3. Performance comparison on text data (accuracy)

t	Reuters			Classic			Sports		
	SL	FedPU	Ext+FedPU	SL	FedPU	Ext+FedPU	SL	FedPU	Ext+FedPU
1%	64.7	65.4	81.7	45.1	45.1	68.5	47.6	72.3	87.8
2%	66.9	67.6	85.1	45.1	45.1	82.2	47.6	78.8	92.9

extension of labeled data by our proposed method (Ext+FedPU). The SL model is trained with only labeled data using the FedAvg method. In Ext+FedPU, the proposed method extends the labeled data and applies the MPUL model [9] with the extended labeled data and the remaining unlabeled data. A linear neural network with an Adam optimizer of learning rate 0.001, batch size 100, and epoch 200 was used in all three compared methods. The extension of labeled data was also performed using a linear neural network with epoch 100, *num_limit* = 200, and *num_add* = 100. The average accuracy in each experimental settings is reported in Table 3. As shown in Table 3, by extending the labeled data using the proposed method, the accuracy of FedPU was greatly improved in all three datasets.

B. Experimental Results using Image Datasets

We evaluated the performance of our model using image data, MNIST, which composed of images with size 2828 from 10 categories, 60,000 training images and 10,000 test images. We set the number of clients as 10 and uniformly divided the data of each class in the training data to clients. In each client, P positive classes with labeled data were randomly chosen and only t percent of data in each positive class was labeled. We tested different settings of P = 1, 2 and t = 1%, 2%. In addition, one more test was performed in the environment where the number of positive classes in 10 clients was set as [1,1,1,1,2,2,2,3,3,3]. All experiments were repeated 10 times with random data splitting for each combination of P and t values.

Neural network models and parameters were set to be the same as those in [9], and the extension of labeled data was performed with *num_limit* = 200 and *num_add* = 100. CNN of two convolution layers (1-3-20 feature maps and 55 kernels) with max-pooling and ReLU function and two fully connected layers (320-50-10) with ReLU function was used. The average accuracy in each experimental settings is reported in Table 4. Ext+FedPU showed improved performance in all cases. However, as the positive classes in clients overlap, the improvement in the performance of Ext+FedPU became smaller.

C. Ablation Study

In the proposed method, two parameters *num_limit* and *num_add* control the number of pseudo-labeled data added to each class. Setting the value of *num_add* as *num_limit/2*, we compared the accuracy in the pseudo-labeled data added by the proposed method by varying

Table 4. Performance comparison on MNIST (accuracy)

P	t	SL	FedPU	Ext+FedPU
[1,1,1,1,1,1,1,1,1,1]	1%	37.7	54.0	61.3
	2%	36.3	65.3	72.6
[2,2,2,2,2,2,2,2,2,2]	1%	36.8	64.9	66.8
	2%	38.6	74.0	76.4
[1,1,1,1,2,2,2,3,3,3]	1%	25.9	62.5	64.7
	2%	24.4	73.5	74.3

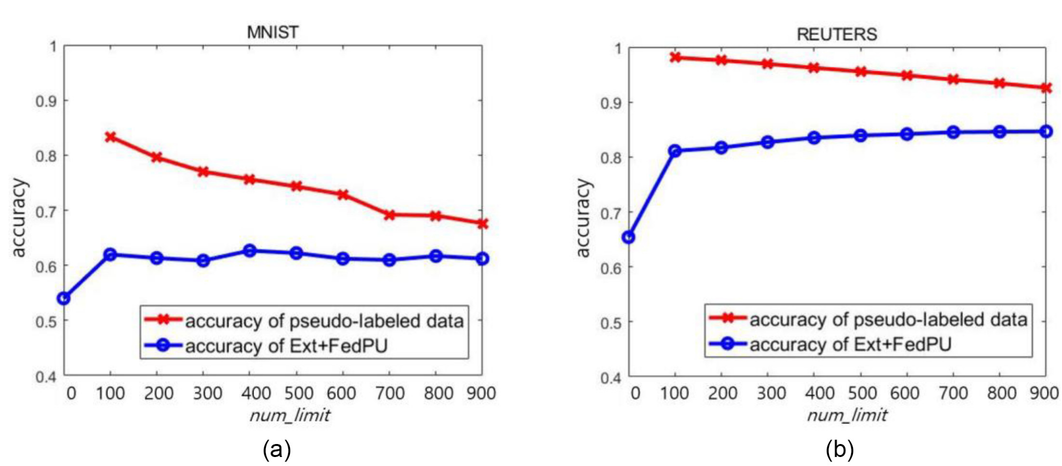

Fig. 4. Comparison of the accuracy in the extended pseudo-labeled data by varying *num_limit*: (a) the MNIST dataset and (b) the Reuters dataset.

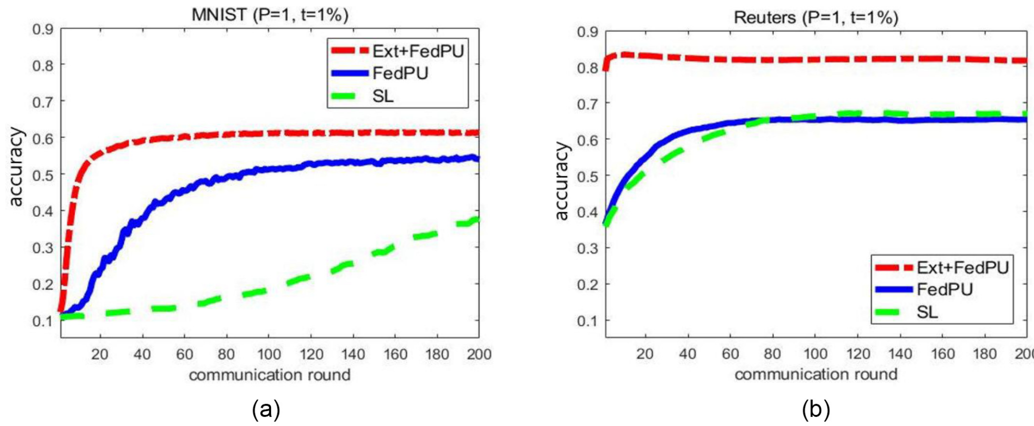

Fig. 5. Comparison of the accuracy in the compared methods according to the communication rounds: (a) the MNIST dataset and (b) the Reuters dataset.

num_limit from 100 to 900. Fig. 4 displays the comparative results in MNIST and Reuters datasets when each client has one positive class (P = 1) and the percentage of labeled data in the positive class is 1% (t = 1). The plot in red shows that the accuracy of pseudo-labeled data is decreasing as the values of *num_limit* on x-axis are increasing. In MNIST data, the accuracy of Ext+FedPU was not changed much; in contrast, in the Reuters data, the accuracy of Ext+FedPU increased regardless of the decrease in the accuracy of the pseudo-labeled data. The results from the Reuters data are believed to be due to the fact that even with *num_limit* set to 900, the accuracy of the pseudo-labeled data is still high, above 0.9.

Fig. 5 shows the accuracy in the compared methods according to the communication rounds. Ext+FedPU achieved high accuracy within short communication rounds, saving the time it takes to reach the optimal global model compared to FedPU or SL methods.

V. CONCLUSION

In this work, we introduce an approach for a federated multi-positive and unlabeled learning method when the size of the labeled data coming from a part of the classes is very small. We adopt a simple PU learning approach for the extension of pseudo-labeled data. In each client, a PU learning model for each positive class is trained by assuming all data except the labeled data of the positive class is negative data. All of the PU learning models are assembled in the server and sent back to each client. Clients extend the set of labeled data using the assembled PU learning models and any semi-supervised federated learning method can be applied since a client has labeled data for all classes. This strategy effectively addresses the challenges posed when small-sized labeled data belongs to a part of the classes in federated learning scenarios. The experimental results, using MNIST data and three text datasets, reveal that our method consistently improves the performance of MPU federated learning method on small-sized multi-positive data and unlabeled data.

CONFLICT OF INTEREST

The author has declared that no competing interests exist.

ACKNOWLEDGMENTS

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No. RS- 2022-00155857, Artificial Intelligence Convergence Innovation Human Resources Development (Chungnam National University)).

REFERENCES

1. C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, "A survey on federated learning, *Knowledge-Based Systems*, vol. 216, article no. 106775, 2021. https://doi.org/10.1016/j.knosys.2021.106775
2. M. Aledhari, R. Razzak, R. M. Parizi, and F. Saeed, "Federated learning: a survey on enabling technologies, protocols, and applications, *IEEE Access*, vol. 8, pp. 140699-140725, 2020. https://doi.org/10.1109/ACCESS.2020.3013541
3. L. Huang, A. L. Shea, H. Qian, A. Masurkar, H. Deng, and D. Liu, "Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records, *Journal of Biomedical Informatics*, vol. 99, article no. 103291, 2019. https://doi.org/10.1016/j.jbi.2019.103291
4. W. Yang, Y. Zhang, K. Ye, L. Li, and C. Z. Xu, "FFD: a federated learning based method for credit card fraud detection,

in *Big Data–BigData 2019*. Cham, Switzerland: Springer, 2019, pp. 18-32. https://doi.org/10.1007/978-3-030-23551-2_2
5. B. Li, Y. Wu, J. Song, R. Lu, T. Li, and L. Zhao, "DeepFed: Federated deep learning for intrusion detection in industrial cyber–physical systems, *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5615-5624, 2021. https://doi.org/10.1109/TII.2020.3023430
6. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data, *Proceedings of Machine Learning Research*, vol. 54, pp. 1273-1282, 2017.
7. T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks, *Proceedings of Machine Learning and Systems*, vol. 2, pp. 429-450, 2020.
8. P. P. Liang, T. Liu, L. Ziyin, N. B. Allen, R. P. Auerbach, D. Brent, R. Salakhutdinov, and L. P. Morency, "Think locally, act globally: federated learning with local and global representations, 2020 [Online]. Available: https://arxiv.org/abs/2001.01523.
9. X. Lin, H. Chen, Y. Xu, C. Xu, X. Gui, Y. Deng, and Y. Wang, "Federated learning with positive and unlabeled data, *Proceedings of Machine Learning Research*, vol. 162, pp. 13344-13355, 2022.
10. J. Bekker and J. Davis, "Learning from positive and unlabeled data: a survey, *Machine Learning*, vol. 109, no. 4, pp. 719-760, 2020. https://doi.org/10.1007/s10994-020-05877-5
11. B. Liu, W. S. Lee, P. S. Yu, and X. Li, "Partially supervised classification of text documents, in *Proceedings of the 19th International Conference (ICML)*, Sydney, Australia, 2002, pp. 387-394. https://dl.acm.org/doi/10.5555/645531.656022
12. J. C. Carnevali, R. G. Rossi, E. Milios, and A. de Andrade Lopes, "A graph-based approach for positive and unlabeled learning, *Information Sciences*, vol. 580, pp. 655-672, 2021. https://doi.org/10.1016/j.ins.2021.08.099
13. M. Du Plessis, G. Niu, and M. Sugiyama, "Convex formulation for learning from positive and unlabeled data, *Proceedings of Machine Learning Research*, vol. 37, pp. 1386-1394, 2015.
14. R. Kiryo, G. Niu, M. C. Du Plessis, and M. Sugiyama, "Positive-unlabeled learning with non-negative risk estimator, *Advances in Neural Information Processing Systems*, vol. 30, pp. 1675-1685, 2017.
15. H. Shi, S. Pan, J. Yang, and C. Gong, "Positive and unlabeled learning via loss decomposition and centroid estimation, *Proceedings of the 27th International Joint Conference on Artificial Intelligence,* Stockholm, Sweden, 2018, pp. 2689-2695. https://dl.acm.org/doi/10.5555/3304889.3305034
16. Y. Xu, C. Xu, C. Xu, and D. Tao, "Multi-positive and unlabeled learning, in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, Melbourne, Australia, 2017, pp. 3182-3188. https://dl.acm.org/doi/10.5555/3172077.3172333
17. S. Shu, Z. Lin, Y. Yan, and L. Li, "Learning from multi-class positive and unlabeled data, in *Proceedings of 2020 IEEE International Conference on Data Mining (ICDM)*, Sorrento, Italy, 2020, pp. 1256-1261. https://doi.org/10.1109/ICDM50108.2020.00160
18. C. H. Park, "Multi-class positive and unlabeled learning for high dimensional data based on outlier detection in a low dimensional embedding space, *Electronics*, vol. 11, no. 17, article no. 2789, 2022. https://doi.org/10.3390/electronics11172789
19. G. Karypis, "Dataset repositories," 2022 [Online]. Available: https://github.com/KarypisLab.

Cheong Hee Park https://orcid.org/0000-0002-8233-2206

Cheong Hee Park received her Ph.D. in Mathematics from Yonsei University, Korea in 1998. She received the M.S. and Ph.D. degrees in Computer Science at the Department of Computer Science and Engineering, University of Minnesota in 2002 and 2004, respectively. She is currently in the dept. of Computer Science and Engineering, Chungnam National University, Korea as a professor. Her research interests include machine learning, data mining, and pattern recognition.