파라미터 변화량을 이용한 연합학습

김한솔, *윤수연
국민대학교
doctorsol@kookmin.ac.kr, 1104py@kookmin.ac.kr

FedAvgP: Federated Averaging with Parameter changes

Hansol Kim, *Sooyeon Yoon
Kookmin Univ.

Abstract

With the development of deep learning and growing interest in personal information protection, federated learning, which learns models with excellent performance without sharing personal data, is attracting attention. However, unlike the training environment, in real scenarios, data heterogeneity problems occur, which causes client drift due to the gap between the global model and the local model. To solve this problem, we propose FedAvgP(Federated Averaging with Parameter changes), a new federated learning parameter aggregation method that uses the amount of change and momentum between the global model and the local model. The proposed method updates global model parameters at the server by utilizing the momentum method and changing the amount between the global model parameters of the previous communication round and the local model parameters of each client in the current round. We demonstrate the performance of the proposed method using computer vision benchmark datasets MNIST, Cifar10, and FMNIST.

Ⅰ. Method

Before explaining the proposed method in detail, we formulate the federated learning environment. We define the number of all clients participating in the federated learning as C, and a specific client as c, $c \in \{1,...,\mathrm{C}\}$. The client learns the local model, and after the training is completed, the number of data used for the training and the parameters of the local model are transferred to the server. Each client has a training dataset $D_{c,N}$. N is the total number of training datasets. we use the momentum method the amount of change between the previous updated global model parameters and the collected local model parameters of each client. The process is as follows:

$$\gamma_c = \frac{\exp(|D_{c,N}||w_g^t - w_c^t|_2)}{\sum_{c=1}^{C}\exp(|D_{c,N}||w_g^t - w_c^t|_2)} \quad (1)$$

$$w_g^{t+1} = \beta w_g^t - \sum_{c=1}^{C}\gamma_c w_c^t \quad (2)$$

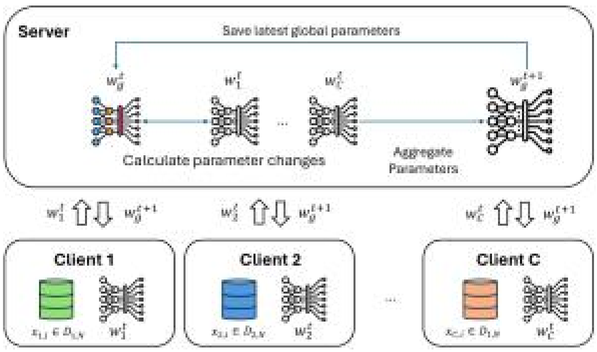

Figure 1.Pipelines of federated average with parameter changes

where t denotes a communication round, $t \in \{0,..., T\}$. w_g denotes the aggregated global model parameter, w_c denotes the local model parameter learned by the client c, $|D_{c,N}|$ denotes the total number of samples of the datasets $D_{c,N}$. β is a momentum coefficient, we set the value to 0.9.

Ⅱ. Experiment

1. Experiment environment

We experiment to prove the performance of the proposed method. Benchmark data use MNIST [4], Cifar10[5], and

*Corresponding author

FMNIST [6], and the data distribution for each benchmark is measured for values of 0.1 and 0.5 based on the Dirichletdistribution. For the performance comparison, classification accuracy on various benchmark data was measured, and the target of performance measurement is representative federated learning methods such as FedAvg [2],FedAvgM[10], FedProx[3] and the proposed method, FedAvgP. Server collects local model parameters and updates the global model parameters, after 5 local epochs. The total communication round is performed 100 times and participation rate of the federated learning is 0.1 and random clients participate in the federated learning. For the experiment, lenet5[11] is used as a backbone model. The optimization function of each client is SGD, and the learning rate is set to 5e-3.

Table 1. Performance comparison table of federated learning methods using computer vision benchmark dataset.

	MNIST		Cifar10		FMNIST	
Method	a=0.1	a=0.5	a=0.1	a=0.5	a=0.1	a=0.5
FedAvg	70.11	88.31	11.05	27.05	68.26	72.21
FedProx	64.08	88.39	17.49	26.67	64.41	72.19
FedAvgM	85.51	94.82	22.72	37.77	69.73	77.2
Ours	90.85	94.91	23.12	38.46	74.94	78.25

2. Classification performance

Among the experimental results, Table 1 shows that the proposed method shows good performance compared to the previous federated learning method. In addition, to further analyze the data classification ability of the model in a data heterogeneous environment, we visualize the result values classified through T-SNE and calculate silhouette-score. The silhouette score is a numerical representation of how closely individual data is clustered with data within its assigned cluster and how far away it is from data in other clusters. A large value can be interpreted as good clustering performance. From Figure 2, we can see that the proposed method classifies data well compared to the previous methods.

Ⅲ. Conclusion

In this paper, we propose FedAvgP (Federated Averaging with Parameter changes), a new aggregate method of federated learning using the parameter changes. The proposed method learns a global model using the amount of change and momentum between the global parameters and the learned parameters of each client. Previously studied federated learning methods had a problem of deteriorating performance in an environment where data was heterogeneous, but the proposed method overcame the limitations and showed excellent performance. We believe that FedAvgP could be applied to various federated learning tasks.

Reference

[1] N. Truong, K. Sun, S. Wang, F. Guitton, and Y. Guo, “Privacy preservation in federated learning: An insightfulsurvey from the gdpr perspective,”

[2] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in Artificial intelligence and statistics. *PMLR, 2017,* pp. 1273–1282.

[3] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, “Federated optimization in heterogeneous networks,” in Proceedings of *Machine Learning and Systems 2020, MLSys 2020, Austin,* TX, USA, March 2-4, 2020, I. S. Dhillon, D. S. Papailiopoulos, and V. Sze, Eds. mlsys.org, 2020.

[4] L. Deng, “The mnist database of handwritten digit images for machine learning research,” *IEEE Signal Processing Magazine,* vol. 29, no. 6, pp. 141–142, 2012.

[5] P. Chrabaszcz, I. Loshchilov, and F. Hutter, “A down-sampled variant of imagenet as an alternative to the cifar datasets,” *arXiv* preprint arXiv:1707.08819, 2017.

[6] H. Xiao, K. Rasul, and R. Vollgraf, “Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms,” *arXiv* preprint arXiv:1708.07747, 2017.

[7] 전주형, and 김중헌. "연합학습 (Federated Learning) 기술동향." *OSIA Standards & Technology Review Journal* 32.2 (2019): 19-22.

[8] Li, Tian, et al. "Federated optimization in heterogeneous networks." *Proceedings of Machine learning and systems 2 (2020)*: 429-450.

[9] Karimireddy, Sai Praneeth, et al. "Scaffold: Stochastic controlled averaging for federated learning." *International conference on machine learning. PMLR, 2020.*

[10] Hsu, Tzu-Ming Harry, Hang Qi, and Matthew Brown. "Measuring the effects of non-identical data distribution for federated visual classification." arXiv preprint arXiv:1909.06335 (2019).

[11] Wei, Guangfen, et al. "Development of a LeNet-5 gas identification CNN structure for electronic noses." *Sensors 19.1 (2019)*: 217.