Federated Learning Paper Sharing

Lisen Dai

November 23, 2020

Goal: reduce the number of communication bits during the models training.

$$\Delta\theta = \mathcal{SGD}_n(\theta, D_{mini-batches}) - \theta$$

θ: Deep Neural Network parameters.
$\mathcal{SGD}_n$: refers to the set of gradient updates after n epochs of SGD on DNN (deep neural network) parameters θ during the sampling of mini-batches from local data
Once we have the updates δv...

FedOpt: Towards Communication Efficiency and Privacy Preservation in Federated Learning

Input: temporal vector $\Delta\theta$, Sparsity Fraction q
Output: sparse temporal $\Delta\theta^*$
Initialization;
$num^+ \leftarrow top_q(\Delta\theta); num^- \leftarrow top_q(-\Delta\theta)$
$\Psi^+ \leftarrow mean(num^+); \Psi^- \leftarrow mean(num^-)$
if $\Psi^+ \geq \Psi^-$ **then**
  return $(\Delta\theta^* \leftarrow \Psi^+(\theta \geq \min(num^+)))$;
end
else
  return $(-\Delta\theta^* \leftarrow \Psi^-(\theta \geq \min(-num^-)))$;
end

Algorithm 1: SCA: Communication Efficiency in FedOpt

FedOpt: Towards Communication Efficiency and Privacy Preservation in Federated Learning

"We utilise the additively homomorphic encryption in FedOpt in order to achieve efficiency throughout the learning process."

Algorithm 2: Pseudocode of Privacy Preserving

Input : Users for local datasets D_i, the cloud server to initialise global parameters ϖ_o
Output: New global parameters ϖ

Initialisation:
while *Cloud server initialise global parameters* ϖ_o **do**
    Aggregate global parameters ϖ_o to users
    while *Users obtain local gradients* G_{II} *by training local models* D_i **do**
        Add noise ϵ-DP $\leftarrow G_{\mathrm{II}}$
        Encrypt $G_{\mathrm{II}} \leftarrow E_\delta(G_{\mathrm{II}} + Lap(\frac{\Delta f_{\mathrm{II}}}{\epsilon}))$
        Generate encrypted local gradients E_{II}
        Aggregate $E_\delta(\sum_{\mathrm{II}=1}^{n} G_{\mathrm{II}})$
    end
    while *Cloud server aggregates encrypted local gradients to users* II **do**
        $E_{add} \leftarrow E_\delta(\sum_{\mathrm{II}=1}^{n} G_{\mathrm{II}})$
        Generate cipher-text from E_{II}
        Generate encrypted global gradients E_{add}
    end
    while *Users decrypts* E_{add} *to get global gradients* B_{II} **do**
        $D_\delta(E_{add}) \leftarrow \sum_{\mathrm{II}=1}^{n} G_{\mathrm{II}}$
        Update existing parameters ϖ
        Aggregate new parameters ϖ to the cloud server
    end
end

FedOpt: Towards Communication Efficiency and Privacy Preservation in Federated Learning

Algorithm 3: FedOpt: Communication-Efficiency and Privacy-Preserving

Input : Initial parameters ϖ

Output: Global model with improved parameters ϖ_o

Initialisation: all users $\Pi_i, i = 1, \ldots,$ [Total number of users] are initialised with the same parameters $v_i \leftarrow v$. Those users who carry different private datasets D_i with $|\{c : (x, y) \in D_i\}|$ = [total classes per user]. The remaining Π are initialised to zero $\Delta v, \mathcal{R}_i, \mathcal{R} \leftarrow 0$.

for *epoch* $e = 1, \ldots, E$ | E = *Total number of Epochs* | **do**

    for $\Pi_i \in \Pi \subseteq \{1, \ldots,$ *[Number of users]*$\}$ **do**

        <u>User Π_i execute:</u>

        Plain-text = $\xi \leftarrow downloads_{CS \rightarrow \Pi_i}(\xi)$

        $\Delta v \leftarrow$ decrypt (ξ)

        $v_i \leftarrow v_i + \Delta v$

        $\Delta v_i \leftarrow \mathcal{R}_i + SGD(v_i, D_i) - v_i$

        $\Delta \overline{v_i} \leftarrow SCA_{upload}(\Delta v_i)$

        $\mathcal{R}_i \leftarrow \Delta v_i - \Delta \overline{v_i}$

        $\xi_i \leftarrow$ encrypt $\Delta \overline{v_i}$

        $upload_{\Pi_i \rightarrow CS}(\xi_i)$

    end

    <u>Cloud Server CS execute:</u>

    $collect_{\Pi_i \rightarrow CS}(\Delta \overline{v_i}), e \in \Pi$

    $\Delta v \leftarrow \mathcal{R} + \frac{1}{\Pi} \sum_{e \in \Pi} \Delta \overline{v_i}$

    $\Delta \overline{v} \leftarrow SCA_{download}(\Delta v)$

    $\mathcal{R} \leftarrow \Delta v - \Delta \overline{v}$

    $v \leftarrow v + \Delta \overline{v_i}$

    $\xi \leftarrow$ encrypt $\Delta \overline{v_i}$

    $Aggregate_{CS \rightarrow \Pi_i}(\xi), i = 1, \ldots, GlobalModel$

end

return ϖ_o

FedOpt: Towards Communication Efficiency and Privacy Preservation in Federated Learning

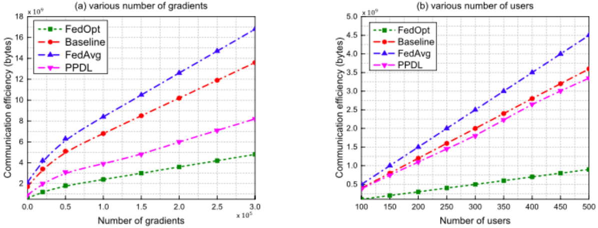

Figure 5. FedOpt communication efficiency on MNIST dataset.

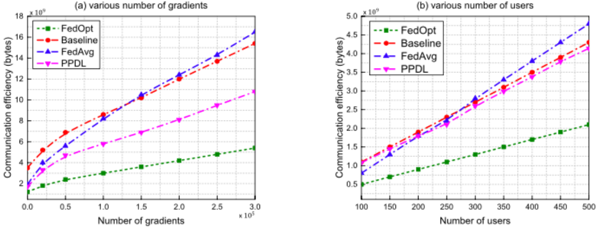

Figure 6. FedOpt communication efficiency on CIFAR-10 dataset.

FedOpt: Towards Communication Efficiency and Privacy Preservation in Federated Learning

Table 2. Communication bits required for upload and download to achieve the targeted accuracy.

	MNIST (Accuracy = 91.3)	**CIFAR-10 (Accuracy = 87.6)**
Baseline	2218/2218 MB	35653 MB/35653 MB
FedAvg $epochs = 50$	119.65 MB/119.65 MB	2589.5 MB/2589.5 MB
FedAvg $epochs = 100$	84.73 MB/84.73 MB	1665.7 MB/1665.7 MB
PPDL $epochs = 50$	98.63 MB/311.6 MB	1472.2 MB/4739.2 MB
PPDL $epochs = 100$	63.74 MB/432.2 MB	958.3 MB/6342.4 MB
FedOpt $epochs = 50$	10.2 MB/102 MB	109.23 MB/1090.3 MB
FedOpt $epochs = 100$	14.6 MB/146 MB	172.3 MB/1723 MB