LaTeX

March 3, 2025

1 Algorithm

Algorithm 1 Process of Generating Data Using Vertical Federated Semi-supervised Method

---

Require: X_{al}^A, X_{nl}^A, $\mathcal{L}_B$, τ, n_{al}, n_{nl}

Ensure: $X^{B_{predict}}$

1: Initialization: $X^{B_{predict}} = \emptyset$, $\mathcal{L}_B^{\text{predict}} = \{(\mu_q, \phi_q^B) \in \mathcal{L}_B \mid u_q > \tau\}$

2: **for** $(\mu_q, \phi_q^B) \in \mathcal{L}_B^{\text{predict}}$ **do**

3:     $X_{al}^{B_{predict}} = \{x_i^{B_{predict}}\}_{i=1}^{n_{al}}$

4:     $X_{nl}^{B_{predict}} = \{x_i^{B_{predict}}\}_{i=n_{al}+1}^{n_{al}+n_{nl}}$

5:     $p = \text{VFPU-M}(X_{al}^A, X_{nl}^A, X_{al}^{B_{predict}}, X_{nl}^{B_{predict}}, \phi_q^B)$

6:     $X^{B_{predict}} = X^{B_{predict}} \cup \{p\}$

7: **end for**

8: **return** $X^{B_{predict}}$

Algorithm 2 VFPU-M Algorithm

Require: X_{al}^A, X_{nl}^A, X_{al}^B, X_{nl}^B, y, α, k, T
Ensure: $\mathbf{y}^{\text{pseudo}}$
1: Initialize the federated model f_θ
2: $\mathcal{D}_L^{(0)} = \{X_{al}^A, X_{al}^B, y\}$
3: $\mathcal{D}_U^{(0)} = \{X_{nl}^A, X_{nl}^B\}$
4: **for** $t = 1$ **to** T **do**
5:     **Federated Model Training:**
6:     $\theta^{(t)} \leftarrow \arg\min_\theta \sum_{(\mathbf{x}^A, \mathbf{x}^B, y) \in \mathcal{D}_L^{(t)}} \ell(f_\theta(\mathbf{x}^A, \mathbf{x}^B), y)$
7:     **Confidence Calculation:**
8:     For $\forall (\mathbf{x}_j^A, \mathbf{x}_j^B) \in \mathcal{D}_U^{(t)}$, calculate the confidence:
9:     $s_j = \begin{cases} \max_c \mathbb{P}(y = c | f_{\theta^{(t)}}(\mathbf{x}_j^A, \mathbf{x}_j^B)) & \text{Classification task} \\ 1 - \frac{|\hat{y}_j - \mu_t|}{\sigma_t} & \text{Regression task} \end{cases}$
10:     Where μ_t, σ_t are the current predicted mean and standard deviation
11:     **Sample Selection:**
12:     $\mathcal{C}^{(t)} = \{(\mathbf{x}_j^A, \mathbf{x}_j^B) \in \mathcal{D}_U^{(t)} | s_j \geq \alpha\}$
13:     Select the top k proportion of samples after sorting by confidence:
14:     $\mathcal{S}^{(t)} = \text{TopK}(\mathcal{C}^{(t)}, k)$
15:     **Pseudo-label Generation:**
16:     $\forall (\mathbf{x}_j^A, \mathbf{x}_j^B) \in \mathcal{S}^{(t)}, \quad \hat{y}_j = \begin{cases} \arg\max_c \mathbb{P}(y = c | f_{\theta^{(t)}}(\mathbf{x}_j^A, \mathbf{x}_j^B)) & \text{Classification} \\ f_{\theta^{(t)}}(\mathbf{x}_j^A, \mathbf{x}_j^B) & \text{Regression} \end{cases}$
17:     **Dataset Update:**
18:     $\mathcal{D}_L^{(t+1)} \leftarrow \mathcal{D}_L^{(t)} \cup \{(\mathbf{x}_j^A, \mathbf{x}_j^B, \hat{y}_j)\}_{j \in \mathcal{S}^{(t)}}$
19:     $\mathcal{D}_U^{(t+1)} \leftarrow \mathcal{D}_U^{(t)} \setminus \mathcal{S}^{(t)}$
20: **end for**
21: $\mathbf{y}^{\text{pseudo}} = \left[y, \bigcup_{t=1}^T \{\hat{y}_j\}_{j \in \mathcal{S}^{(t)}} \right]$ **return** $\mathbf{y}^{\text{pseudo}}$