Federated Multi-Task Learning

General Multi-Task Learning Setup

- General formulation :

$$\min_{W,\Omega} \quad \left[\sum_{k=1}^{K} \sum_{i \in \mathcal{P}_k} f_k(w_k^\top x_k^i, y_k^i) + \mathcal{R}(W, \Omega) \right]$$

$\Omega \in \mathbb{R}^{K \times K}$ models *relationships among tasks*. f_k is a convex loss function.

- MTL problems differ based on their assumptions on $\mathcal{R}$

$$\mathcal{R}(W, \Omega) = \lambda_1 \, tr\left(W \Omega W^\top\right) + \lambda_2 \, | \, W |_F^2$$

Federated Multi-Task Learning

MOCHA : A Framework for Federated Multi-Task Learning

$$\min_{W,\Omega} \quad \left[\sum_{k=1}^{K} \sum_{i \in \mathcal{P}_k} f_k(w_k^\top x_k^i, y_k^i) + \mathcal{R}(W, \Omega) \right]$$

$$\mathcal{R}(W, \Omega) = \lambda_1 \, tr\left(W \Omega W^\top\right) + \lambda_2 \, |W|_F^2$$

- ✓ Not jointly convex in W and Ω
- ✓ When fixing Ω, updating W depends on the data X.
- ✓ When fixing W, optimizing for Ω only depends on W, not on X.

Federated Multi-Task Learning, NIPS, 2017