

# Worker

输入：一个 Batch 数据  $(\mathbf{x}^{k(i)}, \mathbf{y}^{k(i)})$ ，网络模型  $\mathbf{J}$ ，损失函数  $\mathcal{L}$ ，学习率  $\eta$ ，训练轮数  $Epochs$

初始化：  $i = 0$

$i < Epochs$

前向传播：

$$\hat{\mathbf{y}}^{k(i)} = \mathbf{J}(\mathbf{x}^{k(i)}, \mathbf{y}^{k(i)}; \mathbf{W}^k)$$

反向传播：

$$\Delta^{k(i)} = \mathcal{L}(\hat{\mathbf{Y}}^{k(i)}, \mathbf{Y}^{k(i)})$$

$$\mathbf{g}(\mathbf{w}^{k(i)}) = \frac{1}{N} \frac{\partial \Delta^{k(i)}}{\partial \mathbf{w}^{k(i)}}$$

$$\mathbf{g}(\mathbf{b}^{k(i)}) = \frac{1}{N} \frac{\partial \Delta^{k(i)}}{\partial \mathbf{b}^{k(i)}}$$

数据同步：  $\mathbf{g}(\mathbf{w}^{(i)}) = \sum_{k=1}^N \mathbf{g}(\mathbf{w}^{k(i)})$

$$\mathbf{g}(\mathbf{b}^{(i)}) = \sum_{k=1}^N \mathbf{g}(\mathbf{b}^{k(i)})$$

模型更新：根据公式(2-4)  
更新模型参数  $\mathbf{W}$ 、 $\mathbf{B}$

$i++$

输出：最优参数矩阵  $\mathbf{W}$ 、 $\mathbf{B}$