

MT-GB

December 18, 2024

- i : instance index
- r : task index
- $H_{i,r}$: Instance i prediction of the task k
- ch : Common prediction
- rh : Task-specific prediction
- θ_k Task k weight
- $L(y, H_r) = 0.5(y - H_r)^2$: Loss function
- $H_r = (1 - \sigma(\theta_r))ch_i + rh_i$: Obj. func.
- $\frac{\partial L}{\partial ch_i} = -(y - H_r) \cdot (1 - \sigma(\theta_r))$
- $\frac{\partial L}{\partial rh_i} = -(y - H_r)$
- $\frac{\partial L}{\partial \theta_r} = (y - H_r) \cdot \sigma(\theta_r)(1 - \sigma(\theta_r)) \cdot ch_i$