[Ranking $\mathcal{L}$oss Explained]

$$L_{\text{loss}}(\hat{l^p}, l^p) = \max\left(0, -\mathbb{1}(l_i, l_j) \cdot (\hat{l_i} - \hat{l_j}) + \xi\right)$$

$$\text{s.t.} \quad \mathbb{1}(l_i, l_j) = \begin{cases} +1, & \text{if} \quad l_i > l_j \\ -1, & \text{otherwise} \end{cases}$$

① if $l_i > l_j$: $\mathcal{L}(\text{loss}, \hat{l}^p, l^p)$

$= \max(0, -\underbrace{\mathbb{1}(l_i, l_j)}_{+1} \cdot (\hat{l}_i - \hat{l}_j) + \xi)$

$= \max(0, -(\hat{l}_i - \hat{l}_j) + \xi)$

i) if $\hat{l}_i > \hat{l}_j$: correct ← even if correct, $\mathcal{L} > 0$, → minimize $\hat{l}_i, \hat{l}_j$

$\mathcal{L} = \max(0, -\underbrace{(\hat{l}_i - \hat{l}_j)}_{0 < \ \le 1} + \xi)$

→ maximize $\hat{l}_i$, minimize $\hat{l}_j$

ii) if $\hat{l}_i < \hat{l}_j$: wrong

$\mathcal{L} = \max(0, -\underbrace{(\hat{l}_i - \hat{l}_j)}_{-1 \le \ < 0} + \xi)$

$= \max(0, (\hat{l}_j - \hat{l}_i) + \xi)$

→ minimize $\hat{l}_j - \hat{l}_i$

$\boxed{\text{→ maximize } \hat{l}_i, \text{ minimize } \hat{l}_j}$

$$L_{\text{loss}}(\hat{l^p}, l^p) = \max\left(0, -\mathbb{1}(l_i, l_j) \cdot (\hat{l_i} - \hat{l_j}) + \xi\right)$$

$$\text{s.t.} \quad \mathbb{1}(l_i, l_j) = \begin{cases} +1, & \text{if} \quad l_i > l_j \\ -1, & \text{otherwise} \end{cases}$$

2 if $l_i < l_j$:

$$\mathcal{L} = \max(0, -\mathbb{1}(l_i, l_j) \cdot (\hat{l}_i - \hat{l}_j) + \xi)$$
$$= \max(0, (\hat{l}_i - \hat{l}_j) + \xi)$$

i) $\hat{l}_i > \hat{l}_j$: wrong

$$= \max(0, \underbrace{(\hat{l}_i - \hat{l}_j)}_{0 < \quad \leq 1} + \xi)$$

$\boxed{\rightarrow \text{minimize } \hat{l}_i, \text{ maximize } \hat{l}_j}$

ii) $\hat{l}_i < \hat{l}_j$: correct

$$= \max(0, \underset{-1 \leq \quad < 0}{(\hat{l}_i - \hat{l}_j)} + \xi)$$

even if correct, $\mathcal{L} > 0$, $\rightarrow$ minimize $\hat{l}_i, \hat{l}_j$

$\rightarrow$ minimize $\hat{l}_i$, maximize $\hat{l}_j$

$$\frac{1}{B} \sum_{(x,y) \in B^s} \mathcal{L}_{target}(\hat{y}, y) + \lambda \frac{2}{B} \sum_{(x^p, y^p) \in B^s} \mathcal{L}_{loss}(\hat{l}^p, l^p)$$

$$\text{s.t.} \begin{cases} \hat{y} = \Theta_{target}(x) \\ \hat{l}^p = \Theta_{loss}(h^p) \\ l^p = \mathcal{L}_{target}(\hat{y}^p, y^p) \end{cases}$$