The deduction of Binary Alignment Network: Reducing lower bound of domain adaptation via separated alignment

Annoymous Authors

No Institute Given

For source domain D_S, target domain D_T, using transforming function $g\left(\cdot\right)$ to project $\forall x_s \in D_S$, $\forall x_t \in D_T$ to same domain D. The transforming function g satisfies $D_{S'} \cap D_{T'} \neq \oslash$. It can be represented as:

$$\begin{aligned} D_{S'} &= g\left(x_s\right), \\ D_{T'} &= g\left(x_t\right), \\ D &= D_{S'} \cup D_{T'}. \end{aligned}$$

Definition 1.

$$\begin{aligned} \epsilon_{D_{S'}}\left(h\right) &= \epsilon(f_s\left(x_s\right), h\left(x_s\right)), \\ \epsilon_{D_{T'}}\left(h\right) &= \epsilon(f_t\left(x_t\right), h\left(x_t\right)), \\ \lambda &= arg\min_{h \subseteq \mathcal{H}}\left(\epsilon_{D_{S'}}\left(h\right) + \epsilon_{D_{T'}}\left(h\right)\right), \\ d_{\mathcal{H}\Delta\mathcal{H}}\left(D_{S'}, D_{T'}\right) &= 2sup|Pr_{D_{S'}}[I(h)] - Pr_{D_{T'}}[I(h)]|, \end{aligned}$$

Theorem 1. *When* $\epsilon_{D_{T'}}(f_s, f_t) > 0$*, then* $\lambda > 0$*.*

Proof. We suppose $h \in \mathcal{H}$ satisfies $\min\left(\epsilon_{D_{S'}} + \epsilon_{D_{T'}}\right)$, we set the satisfied h as $h^\star$. Such that $\lambda^\star = \left(\epsilon_{D_{S'}}\left(h^\star\right) + \epsilon_{D_{T'}}\left(h^\star\right)\right)$. To simply the equation of λ, we represented as $\lambda = \left(\epsilon_{D_{S'}} + \epsilon_{D_{T'}}\right)$.

$\because$

$$\epsilon_{D_{S'}}\left(h\right), \epsilon_{D_{T'}}\left(h\right) >= 0$$

$\therefore$

$$\begin{aligned} \lambda &> |\epsilon_{D_{S'}} - \epsilon_{D_{T'}}| \\ &\geqslant |\epsilon_{D_{S'}}\left(h^\star, f_s\right) - \epsilon_{D_{T'}}\left(h^\star, f_s\right) - \epsilon_{D_{T'}}\left(f_s, f_t\right)| \end{aligned} \tag{$*$}$$

because

$$d_{\mathcal{H}\Delta\mathcal{H}}\left(D_{S'}, D_{T'}\right) > 0$$

And $d_{\mathcal{H}\Delta\mathcal{H}}\left(D_{S'}, D_{T'}\right)$ can be represented as:

$$d_{\mathcal{H}\Delta\mathcal{H}}\left(D_{S'}, D_{T'}\right) = 2sup \mid \epsilon_{D_{S'}}(h, h') - \epsilon_{T'}(h, h') \mid > 0 \qquad h, h' \in \mathcal{H}$$

We can obviously find a mirror hypothesis $h_{f_s} = f_s, h_{f_s} \in \mathcal{H}$.
s.t.

$$
\begin{aligned}
&\mid \epsilon_{D_{S'}}(h_{f_s}, h') - \epsilon_{T'}(h_{f_s}, h') \mid \\
\leftrightarrow &\mid \epsilon_{D_{S'}}(f_s, h') - \epsilon_{T'}(f_s, h') \mid \geqslant 0
\end{aligned}
$$

Without loss of generality. We assume $k = \mid \epsilon_{D_{S'}}(f_s, h') - \epsilon_{T'}(f_s, h') \mid = 0$. And $(\star)$ can be further represented as :

$$
\mid k - \epsilon_{D_{T'}}(f_s, f_t) \mid \geq \epsilon_{D_{T'}}(f_s, f_t)
$$

■