Learning to cluster.

Transfer learning task:

Source data

$$S = \{x_S, y_S\}$$

target data

$$T = \{x_T, y_T\}$$

$\{x_S\} \neq \{y_S\}$ Across task

transductive learning

$\{x_T\} = \{y_S\}$ with domain shift

$P(x_S) \neq P(x_T)$

labeled source data.

auxiliary data.

$S = S' \cup A$ → may/may not be equal to $\{x_T\}$

$\hookrightarrow \{x_{S'}, y_{S'}\}$

$$\{y_{S'}\} = \{x_T\}$$

for cross task cases: (A & unlabel T

are available

Def: A transfer function $R : R(A) = (x_A^R, y^R)$

$x_A^R = \{(x_{Ai}, x_{Aj})\} \; \forall \; i, j$

{similar, dissimilar}

{ Learn the pairwise function $G(x_i, x_j) = y_{ij}$ } Siamese network

Applying G on T we get

$G(x_{Ti}, x_{Tj}) = y_{Tij}$

infer Y_T from $Y_T^R = \{y_{Tij}\} \; \forall \; i, j$.

Solved by constrained clustering Alg.

Expected ... cluster ...

Learnable clustering Objective

* pair of data x_p, x_q

Their distribution $P = f(x_p)$ } ?? Distribution.

$Q = f(x_q)$

→ Discrete distribution.

similar pair.

$$\ell(x_p, x_q)^+ = D_{KL}(P^* \| Q) + D_{KL}(Q^* \| P)$$

P^*, Q^* assumed to be constant.

dissimilar pair. hinge loss.

$$\ell(x_p, x_q)^- = L_h(D_{KL}(P^* \| Q), \sigma) + L_h(D_{KL}(Q^* \| P), \sigma)$$

→ hinge loss.

$$L_h(e, \sigma) = \max(0, \sigma - e)$$

enforce a margin.

$G(x_p, x_q) \in \{0, 1\}$

$$\mathcal{L}(x_p, x_q) = G(x_p, x_q)\, \mathcal{L}(x_p, x_q)^{+} + \left(1 - G(x_p, x_q)\right) \mathcal{L}(x_p, x_q)^{-}$$

complete loss value: for D set with (p, q)

$$\mathcal{L}_d = \sum_{\forall (p,q) \in D} \mathcal{L}(x_p, x_q)$$

can be combined with other loss.