Learning with Augmented Features for Heterogeneous Domain Adaptation

Key messages

- We present a novel statistical learning method called Heterogeneous Feature Augmentation (HFA) to solve domain adaptation problems with data features of different dimensions, i.e., Heterogeneous Domain Adaptation (HDA).
- The learning of our proposed augmented features can be readily incorporated into different learning methods (e.g., SVM and SVR), making them applicable to HDA tasks.
- Example: Object recognition [Ref 1] and multilingual text categorization [Ref 2]

Background

- Limited or even no labeled training data cannot learn robust classifiers in the domain (i.e., target domain) we are interested in.
- Different domains may have data in different feature spaces (e.g., features of different dimensions).
- Can we learn robust classifiers for the target domain by using existing source domains with their data of a different feature dimension?

Proposed HFA method

Formulation with SVM

\[
\begin{align*}
\min_{P,Q,w,b,\xi^s_k,\xi^t_k} & \frac{1}{2} \|w\|^2 + C \left( \sum_{i=1}^{n_s} \xi^s_k + \sum_{i=1}^{n_t} \xi^t_k \right) \\
\text{s.t.} & \quad y^s_i (w^T x^s_i + b) \geq 1 - \xi^s_k, \xi^s_k \geq 0; \\
& \quad y^t_i (w^T x^t_i + b) \geq 1 - \xi^t_k, \xi^t_k \geq 0; \\
& \quad \|P\|^2 \leq \lambda_p, \|Q\|^2 \leq \lambda_q
\end{align*}
\]

- \( w = [w^s, w^t, w^t]' \) is a feature weight vector
- Taking the dual w.r.t. \( w, b, \xi^s_k \) and \( \xi^t_k \) with the dual variable \( \alpha = [\alpha^s_1, ..., \alpha^s_{n_s}, \alpha^t_1, ..., \alpha^t_{n_t}]' \):

\[
\begin{align*}
\min_{P,Q} & \max_{\alpha} \sum_{i=1}^{n_s} \frac{1}{2} (\alpha \circ y) K_{P,Q} (\alpha \circ y) \\
\text{s.t.} & \quad y^s \alpha = 0, 0_{n_s+n_t} \leq \alpha \leq C 1_{n_s+n_t}; \\
& \quad \|P\|^2 \leq \lambda_p, \|Q\|^2 \leq \lambda_q
\end{align*}
\]

- It is nontrivial to determine the optimal dimension \( d_c \) for the common subspace.
- \( H = [P, Q][P, Q] \in \mathbb{R}^{(d_s+1) \times (d_t+1)} \) free of \( d_c \); \( \tilde{H} \in \mathbb{R}^{(n_s+n_t) \times (n_s+n_t)} \) in kernelized HFA
- Solution to HFA: Iteratively update \( H \) using SDP and \( \alpha \) using SVM

Experiments

Datasets

- Object dataset [Ref 1]: 3 domains + 31 classes
- Reuters multilingual dataset [Ref 2]: 5 domains + 6 classes

Results

<table>
<thead>
<tr>
<th>Source Domain</th>
<th>SVM, T</th>
<th>KCAMC</th>
<th>HeMap</th>
<th>DAMA</th>
<th>Arcs</th>
<th>HFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>web classification</td>
<td>52.9 ± 3.1</td>
<td>50.3 ± 3.0</td>
<td>47.3 ± 2.4</td>
<td>50.3 ± 3.0</td>
<td>47.3 ± 2.4</td>
<td>52.9 ± 3.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source Domain</th>
<th>SVM, T</th>
<th>KCAMC</th>
<th>HeMap</th>
<th>DAMA</th>
<th>Arcs</th>
<th>HFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>web classification</td>
<td>52.9 ± 3.1</td>
<td>50.3 ± 3.0</td>
<td>47.3 ± 2.4</td>
<td>50.3 ± 3.0</td>
<td>47.3 ± 2.4</td>
<td>52.9 ± 3.1</td>
</tr>
</tbody>
</table>

Table 1. Classification accuracy on the object dataset.

Table 2. Classification accuracy on the Reuters multilingual dataset.

References