論文の紹介
InstanceWeighting for Domain Adaptation in NLP
Jing Jiang and ChengXiang Zhai

吉田拓夢

構成

- 全6章

1. Introduction
2. Domain Adaptation
3. A Framework of InstanceWeighting for Domain Adaptation
4. Experiments
5. RelatedWork
6. Conclusions and FutureWork

概要

- インスタンス重み付けによる領域適応問題

1. →導入
2. →領域適応の分析
3. →インスタンス重み付けの手法の提案
4. →実験結果
5. →関連分野との比較
6. →結論

2. Domain Adaptation

- 領域適応問題

$x_i \in X, y_i \in Y$（X:特徴空間　Y:クラスラベル)

$p(y|x;\theta)$←最尤推定でパラメータθの決定

- 分布の異なる2つを扱う

ps(x,y)←ソース(ラベル有)

pt(x,y)←ターゲット(ラベルなし)

→２つのケース

・Labeling Adaptation

・Instance Adaptation

この２つにより
領域適応が
必要とされる

解決策

- Labeling Adaptation

割愛

- Instance Adaptation

→3つの手法

・Ds $\mathcal{D}_s = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$

・Dt,l $\mathcal{D}_{t,l} = \{(x_j^{t,l}, y_j^{t,l})\}_{j=1}^{N_{t,l}}$

・Dt,u $\mathcal{D}_{t,u} = \{x_k^{t,u}\}_{k=1}^{N_{t,u}}$

- Dsを使う

$$
\begin{aligned}
\theta_t^* &\approx \arg\max_{\theta} \int_{\mathcal{X}} \frac{p_t(x)}{p_s(x)} p_s(x) \sum_{y \in \mathcal{Y}} p_s(y|x) \log p(y|x;\theta) dx \\
&\approx \arg\max_{\theta} \int_{\mathcal{X}} \frac{p_t(x)}{p_s(x)} \tilde{p}_s(x) \sum_{y \in \mathcal{Y}} \tilde{p}_s(y|x) \log p(y|x;\theta) dx \\
&= \arg\max_{\theta} \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{p_t(x_i^s)}{p_s(x_i^s)} \log p(y_i^s|x_i^s;\theta).
\end{aligned}
$$

- Dt,lを使う(教師付き学習)

$$
\begin{aligned}
\theta_t^* &\approx \arg\max_\theta \int_{\mathcal{X}} \tilde{p}_{t,l}(x) \sum_{y\in\mathcal{Y}} \tilde{p}_{t,l}(y|x) \log p(y|x;\theta) dx \\
&= \arg\max_\theta \frac{1}{N_{t,l}} \sum_{j=1}^{N_{t,l}} \log p(y_j^{t,l}|x_j^{t,l};\theta)
\end{aligned}
$$

- Dt,uを使う

$$
\begin{aligned}
\theta_t^* &\approx \arg\max_\theta \int_{\mathcal{X}} \tilde{p}_{t,u}(x) \sum_{y\in\mathcal{Y}} p_t(y|x) \log p(y|x;\theta) dx \\
&= \arg\max_\theta \frac{1}{N_{t,u}} \sum_{k=1}^{N_{t,u}} \sum_{y\in\mathcal{Y}} p_t(y|x_k^{t,u}) \log p(y|x_k^{t,u};\theta)
\end{aligned}
$$

3. A Framework of InstanceWeighting for Domain Adaptation

- Labeling AdaptationとInstance Adaptationの２つの解決手法の取り入れ

$$\begin{aligned}\hat{\theta} &= \arg\max_{\theta} \Bigg[\lambda_s \cdot \frac{1}{C_s} \sum_{i=1}^{N_s} \alpha_i \beta_i \log p(y_i^s | x_i^s; \theta) \\ &+ \lambda_{t,l} \cdot \frac{1}{C_{t,l}} \sum_{j=1}^{N_{t,l}} \log p(y_j^{t,l} | x_j^{t,l}; \theta) \\ &+ \lambda_{t,u} \cdot \frac{1}{C_{t,u}} \sum_{k=1}^{N_{t,u}} \sum_{y \in \mathcal{Y}} \gamma_k(y) \log p(y | x_k^{t,u}; \theta) \\ &+ \log p(\theta) \Bigg],\end{aligned}$$

$$C_s = \sum_{i=1}^{N_s} \alpha_i \beta_i.$$

$$C_{t,l} = N_{t,l}$$

$$C_{t,u} = \sum_{k=1}^{N_{t,u}} \sum_{y \in \mathcal{Y}} \gamma_k(y)$$

$$\lambda_s + \lambda_{t,l} + \lambda_{t,u} = 1$$

4. Experiments

- NLPの3つのタスクで実験

・品詞タグ付け　・固有表現分類　・スパムフィルタ

- 設定

・A:ターゲット(少量のラベル有)

1)-ラベル付きのターゲットを利用してソースのラベルを剪定

2)-トレーニングにラベル付きのターゲットを加える

・B:ターゲット(ラベル無し)

1)-ブートストラップ

- A-1),2)の結果

POS		NE Type				Spam			
k	**Oncology**	k	**CTS**	k	**WL**	k	**u00**	**u01**	**u02**
0	0.8630	0	0.7815	0	0.7045	0	0.6306	0.6950	0.7644
4000	0.8675	800	0.8245	600	0.7070	150	0.6417	0.7078	0.7950
8000	0.8709	1600	0.8640	1200	0.6975	300	0.6611	0.7228	0.8222
12000	0.8713	2400	0.8825	1800	0.6830	450	0.7106	0.7806	0.8239
16000	0.8714	3000	0.8825	2400	0.6795	600	0.7911	0.8322	0.8328
all	0.8720	all	0.8830	all	0.6600	all	0.8106	0.8517	0.8067

Table 1: Accuracy on the target domain after removing “misleading” source domain instances.

POS		NE Type			Spam			
method	**Oncology**	**method**	**CTS**	**WL**	**method**	**u00**	**u01**	**u02**
$\mathcal{D}_s$ only	0.8630	$\mathcal{D}_s$ only	0.7815	0.7045	$\mathcal{D}_s$ only	0.6306	0.6950	0.7644
$\mathcal{D}_s + \mathcal{D}_{t,l}$	0.9349	$\mathcal{D}_s + \mathcal{D}_{t,l}$	0.9340	0.7735	$\mathcal{D}_s + \mathcal{D}_{t,l}$	0.9572	0.9572	0.9461
$\mathcal{D}_s + 5\mathcal{D}_{t,l}$	0.9411	$\mathcal{D}_s + 2\mathcal{D}_{t,l}$	0.9355	0.7810	$\mathcal{D}_s + 2\mathcal{D}_{t,l}$	0.9606	0.9600	0.9533
$\mathcal{D}_s + 10\mathcal{D}_{t,l}$	0.9429	$\mathcal{D}_s + 5\mathcal{D}_{t,l}$	0.9360	0.7820	$\mathcal{D}_s + 5D_{t,l}$	0.9628	09611	0.9601
$\mathcal{D}_s + 20\mathcal{D}_{t,l}$	0.9443	$\mathcal{D}_s + 10\mathcal{D}_{t,l}$	0.9355	0.7840	$\mathcal{D}_s + 10\mathcal{D}_{t,l}$	0.9639	0.9628	0.9633
$\mathcal{D}'_s + 20\mathcal{D}_{t,l}$	0.9422	$\mathcal{D}'_s + 10\mathcal{D}_{t,l}$	0.8950	0.6670	$\mathcal{D}'_s + 10\mathcal{D}_{t,l}$	0.9717	0.9478	0.9494

Table 2: Accuracy on the unlabeled target instances after adding the labeled target instances.

- B-1)の結果

	POS	NE Type		Spam		
method	**Oncology**	**CTS**	**WL**	**u00**	**u01**	**u02**
supervised	0.8630	0.7781	0.7351	0.6476	0.6976	0.8068
standard bootstrap	0.8728	0.8917	0.7498	0.8720	0.9212	0.9760
balanced bootstrap	0.8750	0.8923	0.7523	0.8816	0.9256	0.9772

Table 3: Accuracy on the target domain without using labeled target instances. In balanced bootstrapping, more weights are put on the target instances in the objective function than in standard bootstrapping.

6. Conclusions and FutureWork

- 提案したフレームワークは様々な戦略に柔軟
- 教師あり学習>半教師あり学習