

関連ドメインの予測モデルを活用した半教師付きドメイン適合

Semi-supervised Domain Adaptation using Prediction Models in Associated Domains

十河 泰弘 坂井 智哉
Yasuhiro Sogawa Tomoya Sakai

NEC データサイエンス研究所
NEC Corporation

Semi-supervised domain adaptation which trains a prediction model so that it adapts to novel domains from a few labeled and relatively large unlabeled observations. In this talk, we consider semi-supervised domain adaptation and propose a model embedding method. Unlike the conventional semi-supervised domain adaptation, our work utilizes prediction models in source domains. Moreover, our method can generate a pseudo label to unlabeled data without any special assumption on data distribution. Through experiments, we confirm the effectiveness of our proposed model embedding approach.

1. はじめに

人はわずかな事例を観測するだけで、自身が保有する知識や過去の経験から次に何が起こるかを予測することができる。このような特性を機械学習の分野において実現するために、近年ではドメイン適合と呼ばれる手法の研究が盛んに行われている [3, 5, 6]。ドメイン適合は、潤沢なラベル付きサンプルが存在するドメイン（元ドメイン）から得られた知識をラベル付きサンプルが限られたドメイン（目標ドメイン）に適用することで少数サンプルからでも精度が高く予測が可能なモデルを学習することを目的としており、さらにその発展として、目標ドメインでラベル付きサンプルに加えてラベル無しサンプルが存在する場合に、それを活用する半教師付きドメイン適合と呼ばれる手法が提案されている [4]。

本研究では、モデル埋め込みと呼ばれるアプローチに基づいた半教師付きドメイン適合による予測モデルの学習法を提案する。従来のドメイン適合が、元ドメインのデータや元ドメインのモデルパラメータが必要なのにに対し、提案法は元ドメインで学習済みのモデルを用い、そのモデルの特性を表すモデル属性を抽出することでドメイン適合を実現する。また、提案法でラベル無しサンプルへの近似ラベル付与をデータクラスタ性等の仮定なしに実現し、半教師付き学習による予測モデル推定を可能とする。

2. 問題設定

2.1 半教師付きドメイン適合

まず、入力となる d 次元の特徴ベクトルを $\mathbf{x} \in \mathbb{R}^d$ とし、それに対応する出力を $\mathbf{y} \in \mathcal{R}$ とする。元ドメインの数を T とし、予測モデルを構築したい対象ドメインの数を T' と表す。本研究では目標ドメインの $t' \in \{T+1, T+2, \dots, T+T'\}$ に対し、それぞれ $N_{t'}$ 個のラベル付きサンプル $\{(\mathbf{x}_i^{(t')}, y_i^{(t')}) | i = 1, \dots, N_{t'}\}$ が与えられるとする。さらに、これらのラベル付きサンプルに加え、目標ドメインそれぞれについて、 $N'_{t'}$ 個のラベル無しサンプル $\{\mathbf{x}_j | j = 1, \dots, N'\}$ が与えられているとする。一般的な半教師付きドメイン学習では、 T 個の元ドメインで得られるデータまたはモデルのパラメータを基に、 T' 個の目標ド

メインにおける予測モデル $\{f_{t'}\}_{t'=T+1}^{T+T'}$ を構築することである。

2.2 関連ドメインの予測モデル活用

本研究では、従来のドメイン適合と異なり、元ドメインにおいてデータから入出力関係を学習した T 個の予測モデル $\{h_t(\mathbf{x}) | t = 1, \dots, T\}$ が得られていることを仮定する。ここで T は元ドメインの数を、 t は各タスクのインデックスを表し、 $h_t: \mathbb{R}^d \rightarrow \mathbb{R}$ である。これらの予測モデルは任意の学習アルゴリズムによって得られたもので実運用に資するものであり、ある程度の予測精度が保証されているものとする。

本研究での目的は、元ドメインで構築された任意の予測モデル $\{h_t(\mathbf{x}) | t = 1, \dots, T\}$ と目標ドメインのデータ $\{(\mathbf{x}_i^{(t')}, y_i^{(t')}) | i = 1, \dots, N_{t'}\}$ および $\{\mathbf{x}_j | j = 1, \dots, N'\}$ を用いて、目標ドメインにおける予測モデル $\{f_{t'}\}_{t'=T+1}^{T+T'}$ を構築することである。

3. 提案手法

本節では、学習済みモデルを活用した半教師付きドメイン学習を実現するための、モデル埋め込み法を提案する。

3.1 モデル属性とモデル間共通入力表現の獲得

本研究において鍵となるアイディアは、各ドメインでの学習済みモデルの性質を表すモデル属性とそれらに対応するサンプルの変換、モデル間共通入力表現を獲得した上で、目標ドメインにおける予測モデルを構築することである。すなわち、各学習済みモデルを

$$g(\mathbf{z}_i, \mathbf{a}_t) \approx h_t(\mathbf{x}_i)$$

という形で近似することを考える。ここで、 $g: \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}$ 、 $\mathbf{z}_i \in \mathbb{R}^p$ は、 $\mathbf{x}_i$ に対応する p 次元の共通入力表現であり、 $\mathbf{a}_t \in \mathbb{R}^p$ はモデル h_t の特性を表す p 次元のモデル属性ベクトルである。各ドメインに起因した特性は、モデル属性によって表現され、共通入力表現は、モデル属性と合わせて予測を実現するのに用いられる。つまり、ドメイン間で共通した情報は、この共通入力表現が保持する。

まずモデル属性と共通入力表現を推定するために、ラベル付きサンプルとラベル無しサンプルを用いて g と h の近似誤

差を

$$L(\mathbf{Z}, \mathbf{Z}', \mathbf{A}) := L_U(\mathbf{Z}, \mathbf{A}) + \sum_{t'=T+1}^{T+T'} L_L^{(t')}(\mathbf{Z}', \mathbf{A})$$

と定義する。ここで、

$$L_U(\mathbf{Z}, \mathbf{A}) = \sum_{j=1}^{N'} \sum_{t=1}^T \ell(h_t(\mathbf{x}_j), g(\mathbf{z}_j, \mathbf{a}_t)),$$

$$L_L^{(t')}(\mathbf{Z}', \mathbf{A}) = \sum_{i=1}^{N_{t'}} \sum_{t=1}^T \ell(h_t(\mathbf{x}_i^{(t')}), g(\mathbf{z}_i^{(t')}, \mathbf{a}_t))$$

であり、 $\ell: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ は $\ell_S(s, t) = (s - t)^2$ のような任意の損失関数を表し、 $\mathbf{Z} := (\mathbf{z}_1, \dots, \mathbf{z}_{N'})$, $\mathbf{Z}' := (\mathbf{z}_{T+1}, \dots, \mathbf{z}_{T+T'})$, $\mathbf{Z}_t := (\mathbf{z}_1, \dots, \mathbf{z}_{N_{t'}})$, $\mathbf{A} := (\mathbf{a}_1, \dots, \mathbf{a}_T)$. このモデル近似誤差を $\mathbf{z}$ と $\mathbf{a}$ について最小化することで、モデル属性と共通入力表現を推定することができる。また、目標ドメイン t' のモデル属性については、それぞれ上記で得られた共通入力表現を用いて

$$\hat{\mathbf{a}}_{t'} = \arg \min_{\mathbf{a}_{t'}} \sum_{i=1}^{N_{t'}} (y_i^{(t')} - g(\hat{\mathbf{z}}_i^{(t')}, \mathbf{a}_{t'}))^2$$

として求めることができる。

上記の具体例として、 g を下記のモデル属性と共通入力表現の内積で表すことを考える。すなわち、

$$g(\mathbf{z}, \mathbf{a}) = \mathbf{z}^\top \mathbf{a} \quad (1)$$

としたとき、前述のモデル属性及び共通入力表現は、辞書学習 [7] と同様にして推定することができる。

3.2 モデル属性を用いた半教師付きドメイン適合

本節で、推定したモデル属性と共通入力表現を用いた半教師付きドメイン適合による目標ドメインの予測モデル $f_{t'}$ の推定法について述べる。ラベル付きサンプルのみから学習する場合、目標ドメインの予測モデル $f_{t'}$ は、 $f_{t'}$ のモデルパラメータについて下記の経験損失を最小化することで得られる。

$$R_L^{(t')}(f_{t'}) := \sum_{i=1}^{N_{t'}} \ell(y_i^{(t')}, f_{t'}(\mathbf{x}_i^{(t')})).$$

しかしながら、ドメイン適合の問題設定においてはラベル付きサンプルの数が限定されており、正確な予測モデルの学習が難しい。そこで、ラベル無しサンプルの活用を考える。一般的な半教師付き学習 [2] ではラベルなしサンプルへのラベル付与にデータのクラスタ構造等の仮定を導入する必要があるが、本研究では前節で推定したモデル属性を用いてラベル無しサンプルに対し疑似的なラベルを付与することが可能である。結果として、下記の経験損失を最小化することで、目標ドメイン t' の予測モデルを推定することができる。

$$R^{(t')}(f_{t'}) := (1 - \gamma) R_L^{(t')}(f_{t'}) + \gamma R_U(f_{t'}).$$

ここで、

$$R_U(f_{t'}) := \sum_{j=1}^{N'} (g(\hat{\mathbf{z}}_j, \hat{\mathbf{a}}_{t'}) - f_{t'}(\mathbf{x}_j))^2$$

であり、 $g(\hat{\mathbf{z}}_j, \hat{\mathbf{a}}_{t'})$ がラベル無しサンプルに対する疑似ラベルである。また、 γ はラベルありサンプルとラベル無しサンプルのトレードオフパラメータである。

表 1: 100 回試行の平均予測誤差（二乗誤差）。ここで Linear はタスク数 T に依存しないため、各 T に対し共通の値である。

	提案法	Stack	Linear
T=10	18.57	19.54	
T=20	17.50	19.67	19.33
T=40	17.45	21.43	

4. 評価実験

マルチタスク学習 [1] の評価に倣って人工データを生成し、提案手法の評価を行った。本実験では各ドメインのデータ生成に用いる線形モデルの 10 次元重みベクトルを $N(\mathbf{0}_{10}, \Sigma)$, $\Sigma := \text{Diag}\{1, 0.25, 0.1, 0.5, 2.0, 1, 0.25, 0.1, 0.5, 2.0\}$ の 10 次元正規分布から、15 次元の入力特徴ベクトルを $U(-2.5, 2.5)$ の一様分布からランダムに生成した。15 次元の入力特徴ベクトルのうち、ドメイン間で共通の 10 次元が重みベクトルと結びついて出力を与える。比較手法には、リッジ回帰 (Linear) 及びスタッキング (Stack) を用意した。スタッキングのモデルは、 $\min_{\mathbf{w}} \sum_{i=1}^{N_{t'}} (y_i - \sum_{t=1}^T w_t h_t(\mathbf{x}_i))^2$ を解くことで得られる、 $\sum_{t=1}^T w_t h_t(\mathbf{x}_i)$ である。本実験において提案法のパラメータはそれぞれ $p = 5$, $\gamma = 0.1$ とし、モデルの近似に (1) を用いた。また、提案法のモデルおよび元ドメインのモデルにはリッジ回帰モデルを用いた。最終的に 6 個のラベル付きサンプル、200 個のラベル無しサンプルを用いて、学習を行った。

表 1 に 1000 個のテストサンプルに対する各手法の予測誤差を示す。Linear に比べ、提案法がより高精度な予測モデルを学習できていることや、タスク数増加に伴って提案法の精度が改善することがわかる。これらは複数の元ドメインからうまくモデル属性と共通入力表現を獲得して、目標ドメインの学習を行っているためと考えられる。

5. おわりに

本研究では、学習済みモデルを用いたドメイン適合法の提案を行った。提案法により、元ドメインのモデルからモデル属性を抽出することで、半教師付きドメイン適合が可能となり、この有用性を人工データ実験によって確認した。

参考文献

- [1] Andreas Argyriou, Theodoros Evgeniou, and Massimiliano Pontil. Convex multi-task feature learning. *Machine Learning*, Vol. 73, No. 3, pp. 243–272, 2008.
- [2] Olivier Chapelle, Bernhard Schölkopf, and Alexander Zien. *Semi-Supervised Learning*. The MIT Press, 2006.
- [3] Hal Daume III. Frustratingly easy domain adaptation. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pp. 256–263, 2007.
- [4] Hal Daumé III, Abhishek Kumar, and Avishek Saha. Frustratingly easy semi-supervised domain adaptation. In *Proceedings of the 2010 Workshop on Domain Adaptation for Natural Language Processing*, pp. 53–59. Association for Computational Linguistics, 2010.
- [5] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning*, pp. 1126–1135, 2017.
- [6] Raghuraman Gopalan, Ruonan Li, and Rama Chellappa. Domain adaptation for object recognition: An unsupervised approach. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 999–1006. IEEE, 2011.
- [7] Julien Mairal, Francis Bach, and Jean Ponce. Task-driven dictionary learning. *IEEE transactions on pattern analysis and machine intelligence*, Vol. 34, No. 4, pp. 791–804, 2012.