

Нецентрализованная доменная адаптация

.....
.....

Рассматривается задача восстановления совместного распределения по маргинальным.

Ключевые слова:

1 Введение

В последнее время в машинном обучении становятся всё более актуальными задачи такие как zero-shot learning, когда нужно научиться аппроксимировать неизвестное распределение без возможности семплирования из него, и few-shot learning, когда есть небольшой набор семплов неизвестного распределения. В данной работе рассматривается теоретический подход к нахождению совместного распределения по маргинальным, с, возможно, малым размером выборки. Вводится гипотеза о строении совместного распределения всех объектов.

Данная задача решается в пространстве признаков объектов. Также данную задачу можно решать с помощью использования мультимодели, когда мы работаем в пространстве ответов локальных моделей (каждая построена на своем домене).

Определение 1. Под *доменом* понимается подмножество объектов выборки, которые обладают некоторыми одинаковыми признаками.

Когда у нас множество индексов двух доменов не пересекаются, то получается, что объекты этих доменов принадлежат ортогональным подпространствам, то есть проекция распределения одного домена на подпространство другого - функция Дирака и совместное распределение является просто произведением маргинальных. Такая постановка не позволяет найти совместное распределение, поэтому опишем свою гипотезу строения данных.

1.1 Гипотеза об объектах

Считаем, что существует общее пространство объектов (товары в магазине). Есть наблюдаемые параметры этих объектов (для карандашей - цвет грифеля и его мягкость, для книг - объем количество страниц и тип переплета), множество наблюдаемых параметров может как пересекаться (общий параметр для карандашей и книг - цена), так и не пересекаться (цвет грифеля, количество страниц).

1.2 Методы нахождения совместного распределения.

Пусть есть \mathfrak{X} - множество объектов, X_1, X_2 - набор объектов из \mathfrak{X} , множества индексов $\mathcal{A}_1, \mathcal{A}_2 \subset \mathcal{J} = \{1, \dots, n\}$ - множество всех индексов.

- Тогда если $\mathcal{A}_1 \cap \mathcal{A}_2 = \emptyset$ или $X_1 \cap X_2 \neq \emptyset$, то расширяя множество индексов для какого-то домена присваиванием какого-то случайного значения из априорного распределения, будем получать разные результаты совместного распределения (у нас нет никакой общей информации для доменов, поэтому матрица корреляции для совместного распределения будет блочно-диагональной, где блоки - матрицы корреляции маргинальных распределений).
- Если $\mathcal{A}_1 \cap \mathcal{A}_2 \neq \emptyset$ и $X_1 \cap X_2 \neq \emptyset$, тогда аппроксимируя (описывая логистической регрессией) левую часть, можем получить правую:

$$\left\{ \begin{array}{l} p(\mathcal{A}_1 \setminus \mathcal{A}_2 | \mathcal{A}_1 \cap \mathcal{A}_2) \\ p(\mathcal{A}_2 \setminus \mathcal{A}_1 | \mathcal{A}_1 \cap \mathcal{A}_2) \end{array} \right\} \Rightarrow p(\mathcal{A}_1, \mathcal{A}_2 | \mathcal{A}_1 \cap \mathcal{A}_2)$$

2 Постановка задачи аппроксимации параметров квадрата

Дана бинарная картинка:

$$\mathbf{M} \in \{0, 1\}^{m_1 \times m_2}$$

где 1 отвечает черному пикселю картинки, 0 — белому. Введем понятие *изображения* \mathbf{C} - набор координат ненулевых пикселей x_i, y_i :

$$\mathbf{C} \in \mathbb{R}^{N \times 2}$$

где $N = m_1 * m_2$ - число пикселей картинки.

Пусть x_0, y_0 - центр прямоугольника, a, b - размер сторон, тогда прямоугольник можно задать следующими уравнениями:

$$\left| \frac{x - x_0}{a} + \frac{y - y_0}{b} \right| + \left| \frac{x - x_0}{a} - \frac{y - y_0}{b} \right| = 1$$

либо:

$$\left\| \frac{x - x_0}{a}, \frac{y - y_0}{b} \right\|_{\infty} = \max \left(\frac{|x - x_0|}{a}, \frac{|y - y_0|}{b} \right) = 1$$

либо:

$$\left[\left(\frac{x - x_0}{a} \right)^2 - 1 \right] \left[\left(\frac{y - y_0}{b} \right)^2 - 1 \right] = 0$$

Последний вариант - самый простой, для него можно нормально поставить задачу оптимизации, но приобретаем лишние решения.

Литература

- [1] А.А. Адуенко. Выбор мультимodelей в задачах классификации, 2017
http://www.frccsc.ru/sites/default/files/docs/ds/002-073-05/diss/11-adenko/11-Aduenko_main.pdf
- [2] Manuel Pérez-Carrasco and Guillermo Cabrera-Vives and Pavlos Protopapas and Nicolas Astorga and Marouan Belhaj, Adversarial Variational Domain Adaptation, 2019, CoRR
- [3] Garrett Wilson and Diane J. Cook, Adversarial Transfer Learning, 2018, CoRR
- [4] Jing Wang and Jiahong Chen and Jianzhe Lin and Leonid Sigal and Clarence W. de Silva, Discriminative Feature Alignment: Improving Transferability of Unsupervised Domain Adaptation by Gaussian-guided Latent Alignment, 2020
- [5] Jing Jiang, A Literature Survey on Domain Adaptation of Statistical Classifiers, 2008
- [6] А. В. Грабовой, В. В. Стрижов, Анализ выбора априорного распределения для смеси экспертов
- [7] Guo, Jiang and Shah, Darsh J and Barzilay, Regina, Multi-Source Domain Adaptation with Mixture of Experts, 2018, Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, <http://aclweb.org/anthology/D18-1498>