Optimizing Knowledge Transfer Between Domains: A Study of SA and EROT

Moudjahed Mohamed
Sirguey Franck

Abstract

Domain transfer is a crucial task in machine learning, aiming to improve classification performance from a source domain (S) to a target domain (T). This paper focuses on two unsupervised domain transfer approaches: Subspace Alignment (SA) and Entropic Regularized Optimal Transport (EROT). For SA, S and T are represented by subspaces described by eigenvectors. It relies on dimension reduction by principal component analysis (PCA) to extract essential information from each domain, while EROT uses normalization and optimization to transfer points from S to T. The performance of EROT and SA are evaluated on the Office/Caltech dataset in an unsupervised domain transfer context using 1-NN.

1. Introduction

Domain adaptation is an essential problem in machine learning which aims to improve the generalization capacity of a model from one data source to another data source. In many situations, data collected in the source domain may have different characteristics than the target domain, which can result in a significant drop in classification or prediction performance. To address this challenge, various methods have been developed.

This research work focuses on the study of two unsupervised domain adaptation approaches: Subspace Alignment (SA) and Entropically Regularized Optimal Transport (EROT). Subspace Alignment seeks to align the subspaces generated by the source domain and target domain data in a reduced-dimensional space, while Entropically Regularized Optimal Transport is an optimization method aimed at finding the best way of distributing resources from a source domain to a target domain while minimizing costs while respecting specific constraints.

We evaluate these two approaches on the Office/Caltech dataset in an unsupervised domain adaptation context, using a 1-NN classifier. Our goal is to analyze their effectiveness in transferring knowledge between fields. To do this, we explore the impact of various hyperparameters on the performance of the methods and present a detailed evaluation of the results.

This study aims to contribute to a better understanding of domain adaptation techniques and provide useful insights for practical applications of machine learning in scenarios where source and target data exhibit disparities.

2. Use of optimal transport

2.1. Subspace alignment[1]

The "subspace alignment" process aims to align the subspaces generated by the data of S and T in a space of reduced dimension d so as to improve the comparison and classification between the two sets.

Algorithm Overview

- **1 - Dimension reduction with PCA:** The first step consists of reducing the dimension of the S and T data using Principal Component Analysis (PCA). PCA is a data analysis technique that identifies the main directions along which data varies the most. This step makes it possible to reduce the complexity of the data by projecting it into a lower dimensional space d. For this, the PCA function is used to calculate the eigenvectors (principal components) of the data covariance matrix and sort them in descending order based on the corresponding eigenvalues.

- **2 - Calculation of the transformation matrix M:** Once the data S and T have been reduced in dimension, the transformation matrix M is calculated by performing a matrix product between the reduced data of S (X_s) and the reduced data of T (X_t). This matrix makes it possible to capture the relationships between the subspaces of the two sets.

- **3 - Projecting data into aligned space:** The data from S and T are projected into an aligned space using the transformation matrix M. This creates two new datasets S_a and T_a where the subspaces are better aligned.

- **4 - Learning and Prediction :** Finally, a k-nearest neighbors (k-NN) classifier is used to perform data classification with k=1. The 1-NN classifier is trained on the data S_a and the corresponding labels y_s. Once the classifier is trained, it is used to predict the labels of the T_a data.

Algorithm 1 Subspace Alignment DA Algorithm

Require: Data: Source data S, Target data T, Source labels L_S, Subspace dimension d
Ensure: Predicted target labels L_T
$X_S \leftarrow PCA(S, d)$
$X_T \leftarrow PCA(T, d)$
$X_a \leftarrow X_S \cdot X_S^T \cdot X_T$
$S_a \leftarrow S \cdot X_a$
$T_T \leftarrow T \cdot X_T$
$L_T \leftarrow Classify(S_a, T_a, L_S)$

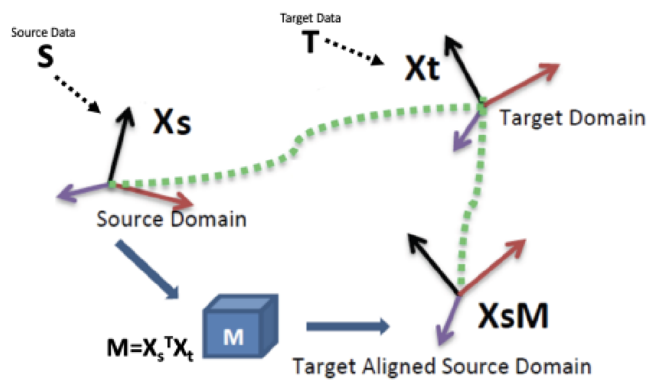

Figure 1. Principle of subspace alignment-based DA

2.2. Entropic regularized optimal transport [3]

Entropic transport regularization is an optimization problem that aims to find the best way to distribute resources from a source to a target while minimizing the overall cost, while taking into account specific constraints. We seek to determine the optimal probability distribution γ which minimizes the objective function as follows[2]:

$$\gamma = \underset{\gamma}{\text{argmin}} \langle \gamma, M \rangle F + \text{reg} \cdot \Omega(\gamma)$$
$$\text{s.t.} \quad \gamma 1 = a, \quad \gamma^T 1 = b, \quad \gamma \geq 0$$

- The metric cost matrix M is of dimension (dim_a, dim_b).
- Ω is the entropic regularization term given by:

$$\Omega(\gamma) = -\sum_{i,j} \gamma_{i,j} \cdot \log(\gamma_{i,j})$$

- a and b are source and target weights (histograms, both sum to 1).

Algorithm Overview

- **1 - Initialization of Distributions:** First of all, we initialize the distributions S and T by assigning them equal and normalized weights. This means that each data point in S and T has the same probability of being selected for transport.
- **2 - Calculation of Cost Matrix:** Next, we calculate a cost matrix M by measuring the distances between each pair of points of S and T using the cdist function. To ensure a fair comparison, we normalize this cost matrix by dividing it by its maximum value.
- **3 - Application of the Sinkhorn Algorithm:** The next step consists of applying the Sinkhorn algorithm using the distributions a and b (initially defined) as well as the cost matrix M and a regularization parameter rege. Sinkhorn's algorithm searches for the best γ transport matrix which minimizes the cost while respecting the constraints imposed by a and b.
- **4 - Readjustment of the Target Distribution:** We calculate the readjusted target distribution S_a by multiplying the transport matrix γ by T. This distribution S_a represents the target distribution modified according to the transport of $S \rightarrow T$.
- **5 - Learning and Prediction:** Finally, we use S_a as training data for a k-nearest neighbors classifier. This classifier is then used to make predictions about T, using the labels associated with the source distribution S.

Algorithm 2 Entropic OT Transport Algorithm

Require: Source data S, Target data T, Regularization parameter $rege$
Ensure: Predicted labels $predictions$
$n_s \leftarrow$ length of S
$n_t \leftarrow$ length of T
$a \leftarrow$ array of size ns with all elements initialized to $\frac{1}{n_s}$
$b \leftarrow$ array of size nt with all elements initialized to $\frac{1}{n_t}$
$M \leftarrow$ Compute cost matrix using cdist(S, T)
$M \leftarrow M/max(M)$
$\gamma \leftarrow$ ot.sinkhorn$(a, b, M, rege)$
$S_a \leftarrow$ Compute transported samples using γ
$L_T \leftarrow Classify(S_a, T, L_S)$

3. Experiments

Method	W→D	D→A	A→C	C→W	W→A
Without DA					
Accuracy	0.31	0.14	0.2	0.12	0.17
SA					
Best d	62	21	78	65	13
Accuracy	0.94	0.34	0.4	0.36	0.37
EROT					
Best rege	0.01	0.02	0.01	0.02	0.04
Accuracy	0.89	0.33	0.31	0.34	0.35

Table 1. Experiments with two domain adaptation methods trained with 1-NN

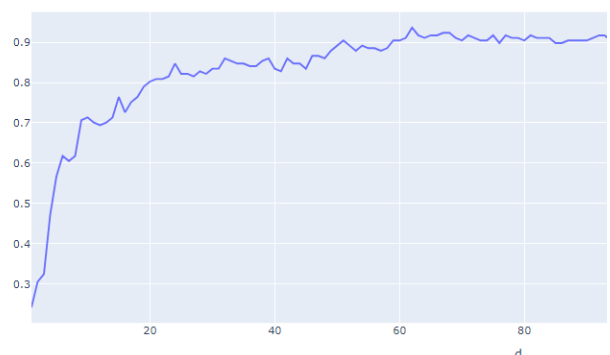

Figure 2. Graph for tuning the parameter d, as a function of accuracy on a 1-NN classifier

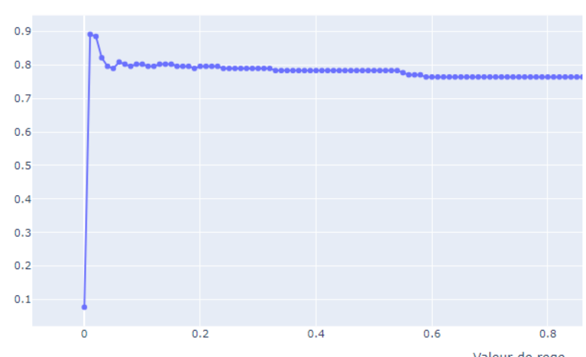

Figure 3. Graph for tuning the parameter rege, as a function of accuracy on a 1-NN classifier

The results of our domain transfer experiments, implementing Subspace Alignment (SA) and Entropic Regularized Optimal Transport (EROT), reveal interesting performance in a variety of domain transfer scenarios. Amazon (A), Caltech (C), Webcam (W), and DSLR (D) data sources were used as source and target domains.

When we look at the performance, we can see that SA generally outperformed EROT in most scenarios. However, it is essential to note that both methods generated significantly better performance than not using the optimal transport. This highlights the positive impact of domain adaptation methods.

A notable result is the transfer from Webcam (W) to DSLR (D), where the SA achieved an impressive 94% accuracy, while the EROT achieved an accuracy of 89%. Without using an optimal domain adaptation method we obtain a precision of 31%, which corresponds to an accuracy approximately multiplied by 3!

4. Conclusion

In summary, our study investigated two unsupervised domain adaptation approaches: Subspace Alignment (SA) and Entropy Regularized Optimal Transport (EROT). We evaluated the performance of these two methods using a 1-NN classifier on the Office/Caltech dataset. Our results demonstrate that, in most scenarios, SA outperforms EROT and that each time domain adaptation significantly increases the accuracy of our predictions, thus illustrating their effectiveness. A particularly notable result was observed when transferring from the Webcam (W) domain to the DSLR (D) domain, where SA achieved an impressive 94% accuracy, while EROT achieved 89% accuracy. These results highlight the positive impact of domain adaptation methods. However, it is important to note that the choice of method depends heavily on the context and specifics of the data. The future of research lies in exploring more sophisticated methods and domain transfer scenarios to improve model generalization.

References

[1] Basura Fernando et al. "Unsupervised Visual Domain Adaptation Using Subspace Alignment". In: (2013).

[2] Rémi Flamary and Nicolas Courty. "API and Modules". In: (2016–2021). URL: https://pythonot.github.io/all.html.

[3] Marc Sebban. "Optimal Transport For Machine Learning". In: (). URL: https://claroline-connect.univ-st-etienne.fr/web/app.php/resource/open/file/1617277.