



ELSEVIER

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Manifold regularized kernel logistic regression for web image annotation



Weifeng Liu^a, Hongli Liu^a, Dapeng Tao^{b,*}, Yanjiang Wang^a, Ke Lu^c

^a China University of Petroleum, China

^b Yunnan University, China

^c University of Chinese Academy of Sciences, China

ARTICLE INFO

Article history:

Received 28 October 2013

Received in revised form

17 April 2014

Accepted 14 June 2014

Available online 9 May 2015

Keywords:

Manifold regularization

Kernel logistic regression

Laplacian eigenmaps

Semi-supervised learning

Image annotation

ABSTRACT

With the rapid advance of Internet technology and smart devices, users often need to manage large amounts of multimedia information using smart devices, such as personal image and video accessing and browsing. These requirements heavily rely on the success of image (video) annotation, and thus large scale image annotation through innovative machine learning methods has attracted intensive attention in recent years. One representative work is support vector machine (SVM). Although it works well in binary classification, SVM has a non-smooth loss function and can not naturally cover multi-class case. In this paper, we propose manifold regularized kernel logistic regression (KLR) for web image annotation. Compared to SVM, KLR has the following advantages: (1) the KLR has a smooth loss function; (2) the KLR produces an explicit estimate of the probability instead of class label; and (3) the KLR can naturally be generalized to the multi-class case. We carefully conduct experiments on MIR FLICKR dataset and demonstrate the effectiveness of manifold regularized kernel logistic regression for image annotation.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Recently, smart devices e.g. smart phone, table PC which equipped with a digital camera have become more and more popular, and people can easily produce millions or even billions of multimedia information such as personal photos or videos. However, it is not convenient to effectively manage the photos or videos at the semantic level, and therefore large scale image/video annotation through innovative machine learning methods has attracted intensive attention in recent years and been successfully deployed for many practical applications in multimedia, computer vision and image processing [14,16,17].

There are a number of machine learning algorithms have been employed for image annotation. One of the representative methods is support vector machine (SVM) that tries to find a separating hyperplane to maximize the margin between two classes [12]. SVM usually minimizes a hinge loss to train the maximum-margin classifier. Although hinge loss is a convex function, it is not

differentiable and can not naturally be generalized to multi-class cases [15].

On the other hand, it is very exhaustive to label a large number of images to learn a robust model for image annotation. Then semi-supervised learning (SSL) has been employed for semi-automatic image annotation [7,8,20]. SSL can improve the generalization ability with only a small number of labeled images by exploiting the intrinsic structure of all the training samples including labeled and unlabeled images [1]. The most traditional class of SSL methods is manifold regularization that tries to explore the geometry of intrinsic data probability distribution by penalizing the objective function along the potential manifold [1,4,9].

Considering the above analysis, in this paper, we replace hinge loss in SVM with logistic loss and propose manifold regularized kernel logistic regression (KLR) for web image annotation. Particularly, we employ the representative Laplacian graph to exploit the geometry of the underlying manifold. Compared to SVM that employs hinge loss function, manifold regularized KLR has the following immediate advantages: (1) the KLR has a smooth loss function; (2) the KLR produces an explicit estimate of the probability instead of class label; (3) the KLR can naturally be generalized to the multi-class case; and (4) Laplacian regularization can well utilize the intrinsic structure of the data distribution.

* Corresponding author at: School of Information Science and Engineering, Yunnan University, Kunming 650091, Yunnan, P.R. China. Tel.: +86 13560476207.

E-mail address: dapeng.tao@gmail.com (D. Tao).

We carefully conduct extensive experiments on the MIR FLICKR dataset. The experimental results verify the effectiveness of Laplacian regularized KLR for web image annotation by comparing it with the baseline algorithms.

The rest of this paper arranged as follows. Section 2 briefly reviews the related work of image classification. Section 3 presents the proposed manifold regularized KLR framework. Section 4 details the implementation of Laplacian regularized KLR. Section 5 demonstrates experimental results on the MIR FLICKR dataset. And Section 6 concludes the paper.

2. Related work

In recent years, there are many algorithms been proposed for multimedia retrieval including image annotation/classification, video indexing, and 3D object retrieval etc. Briefly, the related image/video annotation methods can be divided into three categories based on the employed machine learning schemes which are unsupervised, supervised, and semi-supervised learning.

Unsupervised learning methods use unsupervised machine learning methods such as nonnegative matrix factorization [3], clustering[10] to annotate images/videos.

Supervised learning methods such as support vector machines [13], decision trees[11] aim to find the relationship between labels and visual features. Considering the growing large amount of samples, some active learning methods [14] are introduced to interactively select only effective samples for labeling.

Considering the heavy user labeling effort, semi-supervised learning methods exploit both a small number of labeled samples and a large number of unlabeled samples to boost the generalization of learning model and receive more and more intensive attention recently [9].

3. Manifold regularized kernel logistic regression

In semi-supervised image annotation, we are given a small number of labeled images $S_L = \{(x_i, y_i)\}_{i=1}^l$ and a large number of unlabeled images $S_U = \{x_j\}_{j=l+1}^{l+u}$, where $y_i \in \{+1, -1\}$ is the label of x_i and l, u denote the number of labeled and unlabeled images respectively. Typically, $l \ll u$. Under the assumption of semi-supervised learning, the labeled images $(x, y) \in S_L$ are drawn from a probability P , and unlabeled images $x \in S_U$ are simply drawn from the marginal distribution P_X of P where P_X is a compact manifold M . This assumption indicates that the conditional distribution $P(y|x)$ varies smoothly along the geodesics in the underlying geometry of M and then close images pairs induce similar conditional distribution pairs.

The manifold assumption is widely employed in SSL because it can effectively explore the local geometry of the potential manifold. Then the SSL problem can be written as the following optimization problem by incorporating an additional regularization term to exploit the intrinsic geometry:

$$\min_{f \in H_k} \frac{1}{l} \sum_{i=1}^l \varphi(f, x_i, y_i) + \lambda_1 \|f\|_K^2 + \lambda_2 \|f\|_f^2 \quad (1)$$

where φ is a general loss function, $\|f\|_K^2$ penalizes the classifier complexity in an appropriate reproducing kernel Hilbert space (RKHS) H_k , $\|f\|_f^2$ is the manifold regularization term to penalize f along the underlying manifold, and parameters λ_1 and λ_2 balance the loss function and regularization terms $\|f\|_K^2$ and $\|f\|_f^2$ respectively.

Although there are different choices for the manifold regularization terms $\|f\|_f^2$, Laplacian regularization is promising to preserve the local similarity. In this paper, we introduce the Laplacian

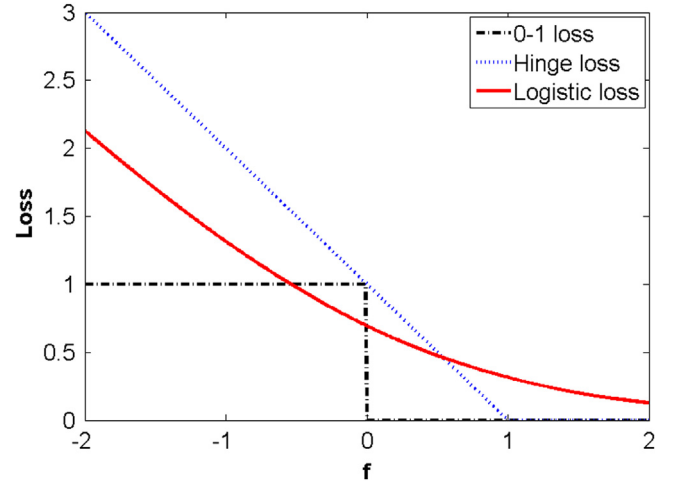


Fig. 1. several loss functions.

regularized kernel logistic regression to web image annotation. In this paper, we employ logistic loss $\log(1 + e^{-f})$ for the loss function to construct a kernel logistic regression (KLR) model. Logistic loss is equivalent to the cross entropy loss function. Some traditional loss functions are plotted in Fig. 1. The dashdot line is 0–1 loss, the dotted line is Hinge loss, and the solid line is logistic loss. From Fig. 1 we can see that the negative log-likelihood loss is smooth and has a similar shape to Hinge loss that used for the SVM. Hence the KLR has similar performance with the SVM. Furthermore, The KLR can naturally be generalized to the multi-class case through kernel multi-logit regression, whereas this is not the case for the SVM [15].

Therefore, we incorporate Laplacian regularized term into the objective function with logistic loss. And then we have the following equivalent optimization problem.

$$\min_{f \in H_k} -\frac{1}{l} \sum_{i=1}^l \left(y_i \log \frac{1}{1 + e^{-f(x_i)}} + (1 - y_i) \log \left(1 - \frac{1}{1 + e^{-f(x_i)}} \right) \right) + \lambda_1 \|f\|_K^2 + \lambda_2 \mathbf{f}^T \mathbf{L} \mathbf{f} \quad (2)$$

where $\mathbf{f} = [f(x_1), f(x_2), \dots, f(x_{l+u})]^T$, L is the graph Laplacian given by $L = D - W$. Here D is a diagonal matrix given by $D_{ii} = \sum_{j=1}^{l+u} W_{ij}$ where W is the edge weight matrix for data adjacency graph.

Theorem 1. The minimization of (2) w.r.t. $f \in H_k$ exists and has the following representation

$$f^* = \sum_{i=1}^{l+u} \theta_i K(x_i, x). \quad (3)$$

where $K(x_i, x)$ is a valid kernel in RKHS.

The theorem 1 is a version of representer theorem [18] that shows the solution of (2) exists and has the general form in terms of the expansion of both labeled and unlabeled images. The proof of this representer theorem can be sketched as below.

Proof: Suppose the subspace $S = \text{span}\{K(x, x_i) | 1 \leq i \leq l+u\}$ is spanned by the kernels centered at labeled and unlabeled images and S^\perp is the orthogonal complement of S . Thus any $f \in H_k$ can be represented as $f = f_S + f_{S^\perp}$, wherein f_S is the projection of f onto S and f_{S^\perp} is the projection of f onto S^\perp . Then we have $\|f\|^2 = \|f_S\|^2 + \|f_{S^\perp}\|^2 \geq \|f_S\|^2$.

On the other hand, K is a valid (symmetric, positive definite) kernel in RKHS and graph Laplacian L is semi-definite positive. Thus $G(\|f\|) = \lambda_1 \|f\|_K^2 + \lambda_2 \|f\|^2 = \lambda_1 \mathbf{f}^T \mathbf{K} \mathbf{f} + \lambda_2 \mathbf{f}^T \mathbf{L} \mathbf{f}$ is a monotonically increasing real-valued function on $\|f\|$. Then we have

Download English Version:

<https://daneshyari.com/en/article/409008>

Download Persian Version:

<https://daneshyari.com/article/409008>

[Daneshyari.com](https://daneshyari.com)