Neural Networks 106 (2018) 96–109

Contents lists available at ScienceDirect

Neural Networks

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

Improved multi-view privileged support vector machine

Jingjing Tang [a], Yingjie Tian [b,c,d,*], Xiaohui Liu [e], Dewei Li [f], Jia Lv [g], Gang Kou [a]

[a] School of Business Administration, Southwestern University of Finance and Economics, Chengdu 611130, China
[b] Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences, Beijing 100190, China
[c] School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190, China
[d] Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, Beijing 100190, China
[e] Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, UK
[f] School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[g] College of Computer and Information Sciences, Chongqing Normal University, Chongqing, 401331, China

HIGHLIGHTS

- IPSVM-MV serves as a general model for multi-view scenario.
- We employ alternating direction method of multipliers to solve IPSVM-MV efficiently.
- We theoretically analyze the performance of IPSVM-MV from two aspects.
- Experimental results demonstrate the effectiveness of the proposed method.

ARTICLE INFO

Article history:
Received 30 December 2017
Received in revised form 24 May 2018
Accepted 29 June 2018

Keywords:
Multi-view learning
Support vector machine
Privileged information
Consensus
Complementarity

ABSTRACT

Multi-view learning (MVL) concentrates on the problem of learning from the data represented by multiple distinct feature sets. The consensus and complementarity principles play key roles in multi-view modeling. By exploiting the consensus principle or the complementarity principle among different views, various successful support vector machine (SVM)-based multi-view learning models have been proposed for performance improvement. Recently, a framework of learning using privileged information (LUPI) has been proposed to model data with complementary information. By bridging connections between the LUPI paradigm and multi-view learning, we have presented a privileged SVM-based two-view classification model, named PSVM-2V, satisfying both principles simultaneously. However, it can be further improved in these three aspects: (1) fully unleash the power of the complementary information among different views; (2) extend to multi-view case; (3) construct a more efficient optimization solver. Therefore, in this paper, we propose an improved privileged SVM-based model for multi-view learning, termed as IPSVM-MV. It directly follows the standard LUPI model to fully utilize the multi-view complementary information; also it is a general model for multi-view scenario, and an alternating direction method of multipliers (ADMM) is employed to solve the corresponding optimization problem efficiently. Further more, we theoretically analyze the performance of IPSVM-MV from the viewpoints of the consensus principle and the generalization error bound. Experimental results on 75 binary data sets demonstrate the effectiveness of the proposed method; here we mainly concentrate on two-view case to compare with state-of-the-art methods.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Real-word data are usually collected from diverse domains or obtained from various feature extractors (see Fig. 1). These data can be naturally partitioned into distinct feature sets, each of which is regarded as a particular view. Thus, each datum can be represented by multiple views and different views often provide information complementary to each other. Multi-view learning (MVL) focuses on learning with multiple views for performance improvement, which is a popular research direction in machine learning field. Compared with single view learning, conventional MVL algorithms are either to concatenate multiple views into one single view with comprehensive description, or to build the learning function for each feature view separately and then jointly optimize the learning

* Corresponding author at: Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences, Beijing 100190, China.
E-mail addresses: tangjingjing13@mails.ucas.ac.cn (J. Tang), tyj@ucas.ac.cn (Y. Tian), xiaohui.liu@brunel.ac.uk (X. Liu), lidewei15@mails.ucas.ac.cn (D. Li), lvjia@cqnu.edu.cn (J. Lv), kougang@swufe.edu.cn (G. Kou).

https://doi.org/10.1016/j.neunet.2018.06.017
0893-6080/© 2018 Elsevier Ltd. All rights reserved.

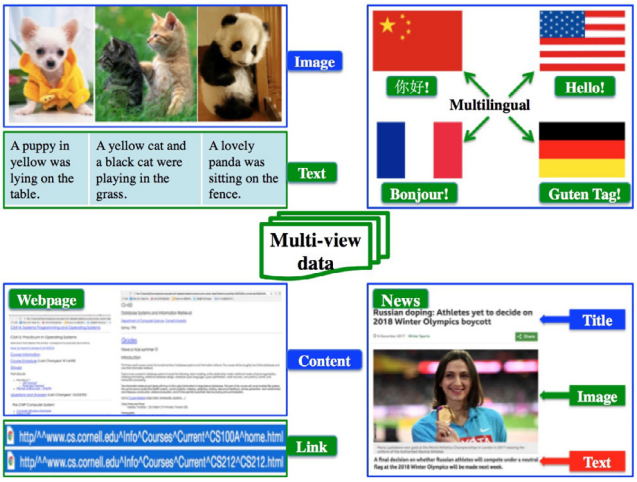

Fig. 1. Multi-view data: (a) an object can be described from different viewpoints such as image and text; (b) multilingual documents have one view in each language; (c) a web document can be represented by text content and citation links on the page; (d) a web news can be depicted by its title, corresponding image and surrounding text.

function by exploiting the redundant views. The concatenating strategy ignores the statistical property of each view and leads to the *curse of dimensionality* problem. While the separation strategy considers each view independently. In fact, views are inherently related since they describe the same set of objects through different feature spaces. A number of methods (Blum & Mitchell, 1998; Xu, Tao, & Xu, 2013) have shown that learning with multiple views jointly is better than the naive approach of using one concatenated view or learning from each view separately.

To date, numerous multi-view learning algorithms have been proposed that can be categorized into three groups (Xu et al., 2013): (1) co-training, (2) multiple kernel learning, and (3) subspace learning. In particular, co-training style algorithms iteratively maximize the mutual agreement on two distinct views to ensure the consistency on the same validation data (Balcan, Blum, & Yang, 2004; Kumar & Daumé, 2011; Li, Nigel, Tao, & Li, 2006; Ménard & Frezza-Buet, 2005; Wang & Zhou, 2010; Wang, Zhang, Wu, Lin, & Zhao, 2017). Multiple kernel learning (MKL) algorithms use kernels that naturally correspond to different views and combine kernels either linearly or non-linearly to improve learning performance (Bach, Lanckriet, & Jordan, 2004; Rakotomamonjy, Bach, Canu, & Grandvalet, 2008; Sonnenburg, Rätsch, Schäfer, & Schölkopf, 2006; Tang & Tian, 2017). Subspace learning algorithms aim to obtain a latent subspace shared by multiple views with the assumption that the input views are generated from this latent subspace (Chao & Sun, 2016; Dhillon, Foster, & Ungar, 2011; Farquhar, Hardoon, Meng, Shawe-taylor, & Szedmak, 2005; Huang, Chung, & Wang, 2016; Liu et al., 2017; Mao & Sun, 2016; Sun & Keates, 2013; Zong, Zhang, Zhao, Yu, & Zhao, 2017). For a comprehensive survey of multi-view learning, please refer to Xu et al. (2013), Sun (2013) and Zhao, Xie, Xu, and Sun (2017).

Although apparent differences exist in the approaches to integrate multiple views for better performance, they mainly embody either *the consensus principle* or *the complementarity principle* (Xu et al., 2013) to ensure their success. The consensus principle aims at maximizing the agreement on the hypotheses among multiple views. For example, SVM-2K (Farquhar et al., 2005) minimizes the distance between the predictive functions of two views as well as the loss within each view by following the consensus principle. In contrast, the complementarity principle emphasizes that each view of the data contains some knowledge not in other views and multiple views share such complementary information to describe the data comprehensively and accurately. The consensus and complementarity principles play key roles in guiding model construction for effective multi-view learning.

From the perspective that MVL targets the complementary knowledge in the learning process, we find another similar framework called learning using privileged information (LUPI) (Lapin, Hein, & Schiele, 2014; Vapnik & Izmailov, 2015; Vapnik & Vashist, 2009). Different from the conventional learning paradigm, where the training data and the test data have the same representations, LUPI models on the training data containing additional information that is only available in the training process. Such additional information is referred as privileged information, and LUPI aims at leveraging the privileged information to boost the performance. A possible analogy of LUPI is human learning with a teacher: when a student learns a geometry concept in school, the teacher acts as the oracle to provide the additional explanations (privileged information) associated with the answers (main information) at any time. However, when later in life the student encounters a new geometry problem, he or she has no access to the teacher's expertise anymore. LUPI incorporates the idea of human teaching for modeling. Under the LUPI paradigm, the standard model SVM_+ (Vapnik & Vashist, 2009) and its reinforced version $\text{SVM}_{\Delta+}$ (Vapnik & Izmailov, 2015) are built on a direct observation that a non-linearly separable (soft-margin) support vector machine can be improved if one had access to a so-called slack oracle (Vapnik & Vashist, 2009). $\text{SVM}_{\Delta+}$ improves SVM_+ by fully utilizing the privileged information to build better classifiers.

LUPI and MVL are similar in the sense that they both try to exploit all the useful information to improve the learning performance, even though there exist some differences between them (as shown in Fig. 2). Considering multiple views and privileged information, the other views can together act as an oracle teacher to give some additional comments for one particular view. That means different views can mutually provide the privileged information to complement and enrich each other. Thus, multiple views share the complementary information analogous to the comments from the teacher in human learning. In this case, it is natural and beneficial to extend LUPI to MVL field.

In our previous work, we have proposed a privileged SVM-based method *PSVM-2V* for two-view learning (Tang, Tian, Zhang, & Liu, 2017). Despite its distinctive advantages, PSVM-2V can be further improved in the following aspects. First, although it utilizes the idea of LUPI to realize the complementarity principle, it just implicitly connects LUPI with the slack variables which are lower bounded by the unknown "correcting function" (explained in Section 3) over the views that are deemed as privileged information. For fully unleashing the power of the complementary information among different feature views, exactly following the LUPI model is beneficial and can bridge better connections between MVL and LUPI. Second, PSVM-2V is built only for two-view learning, which cannot be extended to deal with multi-view (more than two views) problems. Designing a new model for multi-view case is indispensable. Third, PSVM-2V uses the quadratic programming solver for solution. It is certainly preferable to construct a more efficient optimization solver.

In this paper, we propose an improved privileged SVM-based model for multi-view learning, termed as IPSVM-MV. The main contributions of this work are listed as follows:

(1) As a general model for multi-view scenario, IPSVM-MV directly follows the LUPI model to fully utilize the multi-view complementary information.

(2) We employ the alternating direction method of multipliers to obtain the solution of IPSVM-MV efficiently.

Download English Version:

https://daneshyari.com/en/article/6862827

Download Persian Version:

https://daneshyari.com/article/6862827

Daneshyari.com