Neurocomputing 144 (2014) 138–150

Contents lists available at ScienceDirect

Neurocomputing

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

Adaptive multi-view selection for semi-supervised emotion recognition of posts in online student community

Zongkai Yang [a], Zhi Liu [a,*], Sanya Liu [a,b,**], Lei Min [a], Wenting Meng [c]

[a] National Engineering Research Center for E-Learning, Central China Normal University, Wuhan 430079, PR China
[b] Key Laboratory of Adolescent Cyberpsychology and Behavior, Ministry of Education, Wuhan 430079, PR China
[c] Wuhan Branch Company, Hikvision Digital Technology Co., Ltd., Wuhan 430074, PR China

ARTICLE INFO

Article history:
Received 9 November 2013
Received in revised form
7 March 2014
Accepted 9 May 2014
Communicated by Yongdong Zhang
Available online 12 June 2014

Keywords:
Semi-supervised emotion recognition
Multi-view selection
Diversity
Feature emotional strength
Kernel smoothing
Emotion contribution rate

ABSTRACT

In statistical text emotion recognition, semi-supervised learning that can leverage plenty of unlabeled data has drawn much attention in recent years. However, the quality of the training data is typically influenced by some mislabeled samples. In this paper, we present a novel co-training method, namely adaptive multi-view selection (AMVS), to improve labeling accuracy of unlabeled samples for semi-supervised emotion recognition. In particular, two importance distributions are proposed to construct multiple discriminative feature views. One is the distribution of feature emotional strengths, and the other is the importance distribution of view dimensionality. On the basis of these two distributions, several feature views are iteratively selected from the original feature space in a cascaded way, and corresponding base classifiers are trained on these views to build a dynamic and robust ensemble. The experimental results on the real-life dataset consisting of moods posts demonstrate the proposed AMVS outperforms conventional multi-view semi-supervised emotion recognition methods, and that abundant emotional discriminative features could be fully exploited in view selection process.

1. Introduction

With the vigorous development of Web technique, people can freely express personal perspectives and moods anytime through online platforms. On many online learning platforms (i.e., Coursera, Udacity, Edx), the emotional texts are increasingly generated. These texts involve students' hobbies and emotional information on their learning process [1]. In addition, many students like to write the mood posts about their daily lives on the online entertainment platforms such as forum, blog, facebook, microblog, etc. In fact, the emotional information can reflect one's potential psychological conditions. However, openness of the network may lead to some inappropriate online behaviors. In particular, online students are vulnerable to some negative textual information and show radical or depressed moods in writing posts. Under these cases, the ability to recognize sentiments about personal feelings, experience, attitudes towards society etc. is crucial to understand user generated content in student community.

Emotion recognition aims to identify the emotion polarity (i.e., positive or negative) of texts and plays an important role in Web crisis warning. In these years, most of research works on emotion analysis pay attention only to product [2–8] or social reviews [9–15]. Recognizing these reviews with their emotions could provide key summaries to purchasers, which makes it possible to feedback mining on fields in need of improvement and is valuable in business intelligence applications, public opinion monitoring, and personalized recommendation services. However, there are few works focusing on moods posts delivered by college students. Nowadays, many college student users like to indicate their emotional states using several words or strings (e.g., "失去/lose", "鄙视/despise", "!/ah") in virtual community. The possible reason is that these students will inevitably face a series of challenges and problems in campus life, and they prefer to deliver some mood posts online to relieve the pressure. In this study, we dedicate to recognize the emotion (including feelings, moods or attitudes) expressed by students in writing a post.

Currently, supervised learning, which uses complete labeled data to build a robust learner, has achieved significant success in text emotion recognition [2–10]. However, this learning method is not applicable for the practical application scenery in which labeled data are very limited and large amounts of unlabeled data are usually available. Under this situation, some semi-supervised leaning methods, which can utilize large unlabeled data together with limited labeled data, have drawn much attention in emotion

* Corresponding author.
** Corresponding author at: National Engineering Research Center for E-Learning, 152 Luoyu Road, Wuhan, Hubei 430079, PR China. Tel.: +27 67868519; fax: +27 67867024.
E-mail addresses: liuzhi8673@gmail.com (Z. Liu), lsy5918@mail.ccnu.edu.cn (S. Liu).

http://dx.doi.org/10.1016/j.neucom.2014.05.055
0925-2312/

recognition [16–25]. These methods mainly adopt the graph-based learning [16,17], transductive learning [18], hierarchical deep learning [19] or multi-view learning strategy [21–24]. In these methods, multi-view strategy is generally used for co-training in semi-supervised learning. By this method, multiple views can be generated to recognize an object sample in ensemble way. But some related works only use the fixed views dividing way. Actually, much potential emotional information in texts can be helpful for picking large amounts of unlabeled data. So, why not adopt an adaptive multi-view strategy to divide feature space to improve the performance of semi-supervised emotion recognition?

In this paper, we first adopt a language-independent stylometric method [26] which extracts fixed-length *n*-gram Chinese strings as feature set. Thousands of the most frequent *n*-gram strings are used to represent a text. On this basis, an adaptive multi-view selection (AMVS) method is presented to overcome the inadequacies of conventional multi-view methods for semi-supervised emotion recognition. Inspired by the method applied to hyperspectral image classification [27], we introduce two importance distribution into AMVS, i.e., *W* and *D*, denoting the importance distributions of feature emotional strengths and view dimensionality, respectively. To form a view, several features are sampled according to the *W* distribution, and the dimensionality of each view (the number of inner features) is adaptively determined based on *D* distribution. That is, the *W* distribution signifies the discriminating capability all features, and the *D* distribution records the importance of possible views' dimensionalities. Besides, the dimensionality of the view could be dynamically adjusted in each selection process. Finally, the number of views would be determined by a convergence criterion instead of manually setting. Overall, we make the following contributions in this paper:

(1) We proposed a novel multi-view semi-supervised learning method named AMVS for text emotion recognition. AMVS automatically select emotional features to form multiple discriminative views, which gain a more stable ensemble than random sampling.
(2) We introduced the concept of the importance distributions of feature emotional strengths, i.e., *W* distribution. *W* makes full use of emotional information in the feature space, and forms an emotion score for each feature. The emotion scores are used to signify the discriminability of features. The distribution helps to improve the discriminating power of individual views.
(3) Based on the kernel smoothing method, we constructed the importance distribution of view dimensionality, which could be used to determine a suitable dimensionality for each view. Thus, the number of features for each component classifier in ensemble needs not to be the same; in other words, the sizes of the generated views could be different. This is also one of the directions of future work suggested in [28].
(4) The iteration number of generating views could be adaptively controlled with the features' cumulative sampling rate. The strategy effectively reduces the overlapping rate between features in different views, and avoids the generation of redundant views.

The rest of the paper is organized as follows. In Section 2, we introduce some related works. Section 3 describes the general idea of the proposed method. Section 4 introduces the adaptive multi-view selection method for semi-supervised emotion recognition. Experimental settings, results and analysis are reported in Section 5. Section 6 concludes this paper.

2. Related works

Up to now, machine learning based methods have been universally employed for emotion recognition due to their remarkable performance. These methods are based on learning classifiers and utilize many conventional classification algorithms such as support vector machine (SVM) [2,3,5,6], neural networks [8], Naive Bayes [3,14], conditional random field [13] and Bayesian network [7]. An open issue in machine learning based Chinese emotion recognition is the lack of sufficient annotated samples for training, while the supervised learning normally assumes that the sufficient labeled data is available for training an emotion classifier. In fact, it is very difficult and time-consuming to manually label emotional samples. To solve this dilemma, various semi-supervised learning methods are proposed to incorporate unlabeled data to improve the classification performance. Goldberg et al. [16] performed emotion classification experiments on movie reviews by a graph-based semi-supervised learning method. This method adopts a positive emotional sentence percentage based similarity measure method, but the recognition accuracy rate only reaches 59.2% when using 1593 labeled reviews. Sindhwani et al. [17] combined lexical sentiment information, unlabeled data and labeled training data to form a semi-supervised lexical model. This method outperforms purely supervised and competing semi supervised techniques across three domains (enterprise-software products, political election, and online movie reviews). Dasgupta et al. [18] fused various methods in semi-supervised emotion recognition including spectral clustering, active learning, transductive learning, and ensemble learning. The framework effectively distinguishes ambiguous reviews and gains a significant improvement across five domains. Zhou et al. [19] proposed an active deep network (ADN) based on deep belief network and the active learning method. The experimental results on the product reviews demonstrate that ADN is superior to classical semi-supervised recognition methods.

In recent years, much attention has been paid to multi-view style semi-supervised emotion recognition. The method generally divides the feature space into several views (or called feature subspaces) according to the type of language, subjectivity/objectivity of expressions or random sampling rule. Thus, a sample could have multiple views, i.e., there are multiple representations from different feature spaces for a sample. These views could be fused to train an ensemble of classifiers which is usually more robust than a single classifier. Wan [20] applied a co-training mechanism with labeled English corpus and unlabeled Chinese corpus for Chinese emotion classification. Li et al. [21] proposed a personal/impersonal views learning (PIVL) based semi-supervised method, in which each review was partitioned into the two parts, i.e., subjective sentences and objective sentences to train classifiers respectively. They achieved the average accuracy rate of 80.84% which is higher than other single view methods across eight domains. In [22,23], a co-selecting method with dynamic subspace generation (Cotraining-Dynamic) was proposed to iteratively pick informative unlabeled samples, which improved the generalization capability of semi-supervised emotion recognition. Lin et al. [24] developed a co-training algorithm with key sentence extraction (Cotraining-KSE) for semi-supervised emotion recognition on English corpus. In Cotraining-KSE, each review was partitioned into key sentences and details sentences to exploit the key emotional information. The method increased by an average 9.41 percent compared to the benchmark methods. Besides, Ortigosa-Hernández et al. [25] constructed a multi-dimensional classifier which joined the different target variables such as sex, eyes color and hair color information of a person. The method took advantage of the potential statistical information to effectively predict the emotion polarity of large amounts of unlabeled data in