TASC:TOPIC-ADAPTIVE SENTIMENT CLASSIFICATION ON DYNAMIC TWEETS

OBJECTIVE:

The objective of this system is to give sentiment classification based on topic-adaptive for dynamic tweets.

ABSTRACT:

Sentiment classification is a topic-sensitive task, i.e. a classifier trained from one topic will perform worse on another. This is especially a problem for the tweets sentiment analysis. Since the topics in Twitter are very diverse, which makes it impossible to train a universal classifier for all topics? Moreover, compared to product review, Twitter lacks data labeling and rating mechanism to acquire sentiment labels. The extremely sparse text of tweets also brings down the performance of a sentiment classifier. In this paper, we propose a semi-supervised topic-adaptive sentiment classification (TASC) model, which starts with a classifier built on common features and mixed labeled data from various topics. It minimizes the hinge loss to adapt to unlabeled data and features including topic-related sentiment words, authors' sentiments and sentiment connections derived from "@" mentions of tweets, named as topic-adaptive features. Text and non-text features are extracted and naturally split into two views for co-training. The TASC learning algorithm updates topic-adaptive features based on the collaborative selection of unlabeled data, which in turn helps to select more reliable tweets to boost the performance. We also design the adapting model along a timeline (TASC-t) for dynamic tweets. Experiment on 6 topics from published tweet corpuses demonstrates that TASC outperforms other well-known supervised and ensemble classifiers. It also beats those semi-supervised learning methods without feature adaption. Meanwhile TASC-t for dynamic tweets can also achieve impressive accuracy and F-score. Finally with timeline visualization of "river" graph, people can intuitively grasp the ups and downs of sentiments' evolvement, and the intensity by color gradation.

INTRODUCTION:

Topics discussed in Twitter are more diverse and un-predictable. Sentiment classifiers always dedicate them-selves to a specific domain or topic named in the paper. Namely, a classifier trained on sentiment data from one topic often performs poorly on test data from another. One of the main reasons is that words and even language constructs used for expressing sentiments can be quite different on different topics. Taking a comment "read the book" as an example, it could be positive in a book review while negative in a movie review. In social media, a Twitter user may have different opinions on different topics. Thus, topic adaptation is needed for sentiment classification of tweets on emerging and unpredictable topics.

EXISTING SYSTEM

 In the existing system based on the initial expansion of the   words they going to give sentiment process based on topics they adopted. Initially in the existing system iteration process is done (for example 1st iteration they are going to take 100 tweets, within that 100 tweets which words are coming with more positive or more negative count that words will be added as positive or negative before 2nd iteration).Here accuracy is less because after iteration immediately we considering positive or negative sentiment without considering left out words in tweets.