Joint Modeling of News Reader's and Comment Writer's Emotions

Huanhuan Liu[†] **Shoushan Li**[†‡*] **Guodong Zhou**[†] **Chu-Ren Huang**[‡] **Peifeng Li**[†]

[†]Natural Language Processing Lab
Soochow University, China
{huanhuanliu.suda,shoushan.li,
churenhuang}@gmail.com

[‡]Department of CBS
the Hong Kong Polytechnic University
{gdzhou,pfli}@suda.edu.cn

Abstract

Emotion classification can be generally done from both the writer's and reader's perspectives. In this study, we find that two foundational tasks in emotion classification, i.e., reader's emotion classification on the news and writer's emotion classification on the comments, are strongly related to each other in terms of coarse-grained emotion categories, i.e., *negative* and *positive*. On the basis, we propose a respective way to jointly model these two tasks. In particular, a co-training algorithm is proposed to improve semi-supervised learning of the two tasks. Experimental evaluation shows the effectiveness of our joint modeling approach.

1 Introduction

Emotion classification aims to predict the emotion categories (e.g., *happy*, *angry*, or *sad*) of a given text (Quan and Ren, 2009; Das and Bandyopadhyay, 2009). With the rapid growth of computer mediated communication applications, such as social websites and miro-blogs, the research on emotion classification has been attracting more and more attentions recently from the natural language processing (NLP) community (Chen et al., 2010; Purver and Battersby, 2012).

In general, a single text may possess two kinds of emotions, writer's emotion and reader's emotion, where the former concerns the emotion expressed by the writer when writing the text and the latter concerns the emotion expressed by a reader after reading the text. For example, consider two short texts drawn from a news and corresponding comments, as shown in Figure 1. On

one hand, for the news text, while its writer just objectively reports the news and thus does not express his emotion in the text, a reader could yield *sad* or *worried* emotion. On the other hand, for the comment text, its writer clearly expresses his *sad* emotion while the emotion of a reader after reading the comments is not clear (Some may feel *sorry* but others might feel *careless*).

News:
> Today's Japan earthquake could be
> 2011 quake aftershock. ......

News Writer's emotion: None
News Reader's emotion: *sad*, *worried*

Comments:
> (1) *I hope everything is ok, so sad. I still can not forget last year.*
> (2) *My father-in-law got to experience this quake... what a suffering.*

Comment Writer's emotion: *sad*
Comment Reader's emotion: Unknown

Figure 1: An example of writer's and reader's emotions on a news and its comments

Accordingly, emotion classification can be grouped into two categories: reader's emotion and writer's emotion classifications. Although both emotion classification tasks have been widely studied in recent years, they are always considered independently and treated separately.

However, news and their corresponding comments often appear simultaneously. For example, in many news websites, it is popular to see a news followed by many comments. In this case, because the writers of the comments are a part of the readers of the news, the writer's emotions on the comments are exactly certain reflection of the reader's emotions on the news. That is, the comment writer's emotions and the news reader's emotions are strongly related. For example,

---

* Corresponding author

in Figure 1, the comment writer's emotion 'sad' is among the news reader's emotions.

Above observation motivates joint modeling of news reader's and comment writer's emotions. In this study, we systematically investigate the relationship between the news reader's emotions and the comment writer's emotions. Specifically, we manually analyze their agreement in a corpus collected from a news website. It is interesting to find that such agreement only applies to coarse-grained emotion categories (i.e., *positive* and *negative*) with a high probability and does not apply to fine-grained emotion categories (e.g., *happy*, *angry*, and *sad*). This motivates our joint modeling in terms of the coarse-grained emotion categories. Specifically, we consider the news text and the comment text as two different views of expressing either the news reader's or comment writer's emotions. Given the two views, a co-training algorithm is proposed to perform semi-supervised emotion classification so that the information in the unlabeled data can be exploited to improve the classification performance.

2 Related Work

2.1 Comment Writer's Emotion Classification

Comment writer's emotion classification has been a hot research topic in NLP during the last decade (Pang et al., 2002; Turney, 2002; Alm et al., 2005; Wilson et al., 2009) and previous studies can be mainly grouped into two categories: coarse-grained and fine-grained emotion classification.

Coarse-grained emotion classification, also called sentiment classification, concerns only two emotion categories, such as *like* or *dislike* and *positive* or *negative* (Pang and Lee, 2008; Liu, 2012). This kind of emotion classification has attracted much attention since the pioneer work by Pang et al. (2002) in the NLP community due to its wide applications (Cui et al., 2006; Riloff et al., 2006; Dasgupta and Ng, 2009; Li et al., 2010; Li et al., 2011).

In comparison, fine-grained emotion classification aims to classify a text into multiple emotion categories, such as *happy*, *angry*, and *sad*. One main group of related studies on this task is about emotion resource construction, such as emotion lexicon building (Xu et al., 2010; Volkova et al., 2012) and sentence-level or document-level corpus construction (Quan and Ren, 2009; Das and Bandyopadhyay, 2009). Besides, all the related studies focus on supervised learning (Alm et al., 2005; Aman and Szpakowicz, 2008; Chen et al., 2010; Purver and Battersby, 2012; Moshfeghi et al., 2011), and so far, we have not seen any studies on semi-supervised learning on fine-grained emotion classification.

2.2 News Reader's Emotion Classification

While comment writer's emotion classification has been extensively studied, there are only a few studies on news reader's emotion classification from the NLP and related communities.

Lin et al. (2007) first describe the task of reader's emotion classification on the news articles and then employ some standard machine learning approaches to train a classifier for determining the reader's emotion towards a news. Their further study, Lin et al. (2008) exploit more features and achieve a higher performance.

Unlike all the studies mentioned above, our study is the first attempt on exploring the relationship between comment writer's emotion classification and news reader's emotion classification.

3 Relationship between News Reader's and Comment Writer's Emotions

To investigate the relationship between news reader's and comment writer's emotions, we collect a corpus of Chinese news articles and their corresponding comments from Yahoo! Kimo News (http://tw.news.yahoo.com), where each news article is voted with emotion tags from eight categories: *happy*, *sad*, *angry*, *meaningless*, *boring*, *heartwarming*, *worried*, and *useful*. These emotion tags on each news are selected by the readers of the news. Note that because the categories of "*useful*" and "*meaningless*" are not real emotion categories, we ignore them in our study. Same as previous studies of Lin et al. (2007) and Lin et al. (2008), we consider the voted emotions as reader's emotions on the news, i.e., the news reader's emotions. We only select the news articles with a dominant emotion (possessing more than 50% votes) in our data. Besides, as we attempt to consider the comment writer's emotions, the news articles without any comments are filtered.

As a result, we obtain a corpus of 3495 news articles together with their comments and the numbers of the articles of *happy*, *sad*, *angry*, *boring*, *heartwarming*, and *worried* are 1405, 230, 1673, 75, 92 and 20 respectively. For coarse-grained categories, *happy* and *heartwarming* are merged into the *positive* category while

sad, *angry, boring* and *worried* are merged into the *negative* category.

Besides the tags of the reader's emotions, each news article is followed by some comments, which can be seen as a reflection of the writer's emotions (Averagely, each news is followed by 15 comments). In order to know the exact relationship between these two kinds of emotions, we select 20 news from each category and ask two human annotators, named **A** and **B**, to manually annotate the writer's emotion (single-label) according to the comments of each news. Table 1 reports the agreement on annotators and emotions, measured with Cohen's kappa (κ) value (Cohen, 1960).

	κ Value (Fine-grained emotions)	κ Value (Coarse-grained emotions)
Annotators	0.566	0.742
Emotions	0.504	0.756

Table 1: Agreement on annotators and emotions

Agreement between two annotators: The annotation agreement between the two annotators is 0.566 on the fine-grained emotion categories and 0.742 on the coarse-grained emotion categories.

Agreement between news reader's and comment writer's emotions: We compare the news reader's emotion (automatically extracted from the web page) and the comment writer's emotion (manually annotated by annotator **A**). The annotation agreement between the two kinds of emotions is 0.504 on the fine-grained emotion categories and 0.756 on the coarse-grained emotion categories. From the results, we can see that the agreement on the fine-grained emotions is a bit low while the agreement between the coarse-grained emotions, i.e., *positive* and *negative*, is very high. We find that although some fine-grained emotions of the comments are not consistent with the dominant emotion of the news, they belong to the same coarse-grained category.

In a word, the agreement between news reader's and comment writer's emotions on the coarse-grained emotions is very high, even higher than the agreement between the two annotators (0.754 vs. 0.742).

In the following, we focus on the coarse-grained emotions in emotion classification.

4 Joint Modeling of News Reader's and Comment Writer's Emotions

Given the importance of both news reader's and comment writer's emotion classification as de-scribed in Introduction and the close relationship between news reader's and comment writer's emotions as described in last section, we systematically explore their joint modeling on the two kinds of emotion classification.

In semi-supervised learning, the unlabeled data is exploited to improve the models with a small amount of the labeled data. In our approach, we consider the news text and the comment text as two different views to express the news or comment emotion and build the two classifiers C_N and C_C. Given the two-view classifiers, we perform co-training for semi-supervised emotion classification, as shown in Figure 2, on both news reader's and comment writer's emotion classification.

Input:

L_{News} the labeled data on the news

$L_{Comment}$ the labeled data on the comments

U_{News} the unlabeled data on the news

$U_{Comment}$ the labeled data on the comments

Output:

L_{News} New labeled data on the news

$L_{Comment}$ New labeled data on the comments

Procedure:

Loop for N iterations until $U_{News} = \varnothing$ or $U_{Comment} = \varnothing$

(1). Learn classifier C_N with L_{News}

(2). Use C_N to label the samples from U_{News}

(3). Choose n_1 *positive* and n_1 *negative* news N_1 most confidently predicted by C_N

(4). Choose corresponding comments M_1 (the comments of the news in N_1)

(5). Learn classifier C_C with $L_{Comment}$

(6). Use C_C to label the samples from $U_{Comment}$

(7). Choose n_2 *positive* and n_2 *negative* comments M_2 most confidently predicted by C_C

(8). Choose corresponding comments N_2 (the news of the comments in M_2)

(9). $L_{News} = L_{News} + N_1 + N_2$

$L_{Comment} = L_{Comment} + M_1 + M_2$

(10). $U_{News} = U_{News} - N_1 - N_2$

$U_{Comment} = U_{Comment} - M_1 - M_2$

Figure 2: Co-training algorithm for semi-supervised emotion classification

5   Experimentation

5.1   Experimental Settings

Data Setting: The data set includes 3495 news articles (1572 positive and 1923 negative) and their comments as described in Section 3. Although the emotions of the comments are not given in the website, we just set their coarse-grained emotion categories the same as the emotions of their source news due to their close relationship, as described in Section 3. To make the data balanced, we randomly select 1500 positive and 1500 negative news with their comments for the empirical study. Among them, we randomly select 400 news with their comments as the test data.

   Features: Each news or comment text is treated as a bag-of-words and transformed into a binary vector encoding the presence or absence of word unigrams.

   Classification algorithm: the maximum entropy (ME) classifier implemented with the public tool, Mallet Toolkits[*].

5.2   Experimental Results

News reader's emotion classifier: The classifier trained with the news text.

   Comment writer's emotion classifier: The classifier trained with the comment text.

   Figure 3 demonstrates the performances of the news reader's and comment writer's emotion classifiers trained with the 10 and 50 initial labeled samples plus automatically labeled data from co-training. Here, in each iteration, we pick 2 positive and 2 negative most confident samples, i.e., $n_1 = n_2 = 2$. From this figure, we can see that our co-training algorithm is very effective: using only 10 labeled samples in each category achieves a very promising performance on either news reader's or comment writer's emotion classification. Especially, the performance when using only 10 labeled samples is comparable to that when using more than 1200 labeled samples on supervised learning of comment writer's emotion classification.

   For comparison, we also implement a self-training algorithm for the news reader's and comment writer's emotion classifiers, each of which automatically labels the samples from the unlabeled data independently. For news reader's emotion classification, the performances of self-training are 0.783 and 0.79 when 10 and 50 ini-

---

tial labeled samples are used. For comment writer's emotion classification, the performances of self-training are 0.505 and 0.508. These results are much lower than the performances of our co-training approach, especially on the comment writer's emotion classification i.e., 0.505 and 0.508 vs. 0.783 and 0.805.
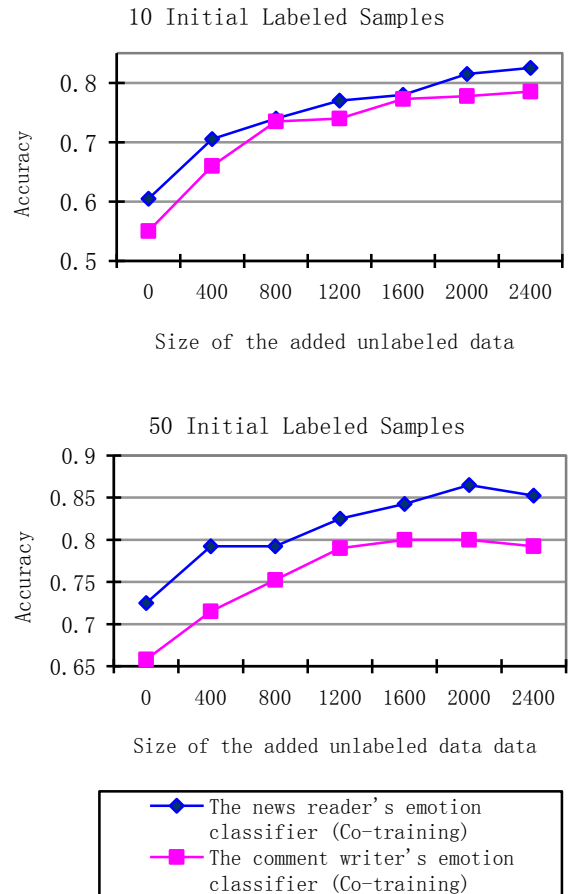

Figure 3: Performances of the news reader's and comment writer's emotion classifiers using the co-training algorithm

6   Conclusion

In this paper, we focus on two popular emotion classification tasks, i.e., reader's emotion classification on the news and writer's emotion classification on the comments. From the data analysis, we find that the news reader's and comment writer's emotions are highly consistent to each other in terms of the coarse-grained emotion categories, *positive* and *negative*. On the basis, we propose a co-training approach to perform semi-supervised learning on the two tasks. Evaluation shows that the co-training approach is so effective that using only 10 labeled samples achieves nice performances on both news reader's and comment writer's emotion classification.

References

Alm C., D. Roth and R. Sproat. 2005. Emotions from Text: Machine Learning for Text-based Emotion Prediction. In *Proceedings of EMNLP-05*, pp.579-586.

Aman S. and S. Szpakowicz. 2008. Using Roget's Thesaurus for Fine-grained Emotion Recognition. In *Proceedings of IJCNLP-08*, pp.312-318.

Chen Y., S. Lee, S. Li and C. Huang. 2010. Emotion Cause Detection with Linguistic Constructions. In *Proceeding of COLING-10*, pp.179-187.

Cohen J. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1):37–46.

Cui H., V. Mittal and M. Datar. 2006. Comparative Experiments on Sentiment Classification for Online Product Comments. In *Proceedings of AAAI-06*, pp.1265-1270.

Das D. and S. Bandyopadhyay. 2009. Word to Sentence Level Emotion Tagging for Bengali Blogs. In *Proceedings of ACL-09*, pp.149-152.

Dasgupta S. and V. Ng. 2009. Mine the Easy, Classify the Hard: A Semi-Supervised Approach to Automatic Sentiment Classification. In *Proceedings of ACL-IJCNLP-09*, pp.701-709, 2009.

Duin R. 2002. The Combining Classifier: To Train Or Not To Train? In *Proceedings of 16th International Conference on Pattern Recognition (ICPR-02)*.

Fumera G. and F. Roli. 2005. A Theoretical and Experimental Analysis of Linear Combiners for Multiple Classifier Systems. *IEEE Trans. PAMI*, vol.27, pp.942–956, 2005.

Li S., Z. Wang, G. Zhou and S. Lee. 2011. Semi-supervised Learning for Imbalanced Sentiment Classification. In *Proceeding of IJCAI-11*, pp.826-1831.

Li S., C. Huang, G. Zhou and S. Lee. 2010. Employing Personal/Impersonal Views in Supervised and Semi-supervised Sentiment Classification. In *Proceedings of ACL-10*, pp.414-423.

Lin K., C. Yang and H. Chen. 2007. What Emotions do News Articles Trigger in Their Readers? In *Proceeding of SIGIR-07*, poster, pp.733-734.

Lin K., C. Yang and H. Chen. 2008. Emotion Classification of Online News Articles from the Reader's Perspective. In *Proceeding of the International Conference on Web Intelligence and Intelligent Agent Technology*, pp.220-226.

Liu B. 2012. *Sentiment Analysis and Opinion Mining (Introduction and Survey)*. Morgan & Claypool Publishers, May 2012.

Kittler J., M. Hatef, R. Duin, and J. Matas. 1998. On Combining Classifiers. *IEEE Trans. PAMI*, vol.20, pp.226-239, 1998

Moshfeghi Y., B. Piwowarski and J. Jose. 2011. Handling Data Sparsity in Collaborative Filtering using Emotion and Semantic Based Features. In *Proceedings of SIGIR-11*, pp.625-634.

Pang B. and L. Lee. 2008. Opinion Mining and Sentiment Analysis: Foundations and Trends. *Information Retrieval*, vol.2(12), 1-135.

Pang B., L. Lee and S. Vaithyanathan. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. In *Proceedings of EMNLP-02*, pp.79-86.

Purver M. and S. Battersby. 2012. Experimenting with Distant Supervision for Emotion Classification. In *Proceedings of EACL-12*, pp.482-491.

Quan C. and F. Ren. 2009. Construction of a Blog Emotion Corpus for Chinese Emotional Expression Analysis. In *Proceedings of EMNLP-09*, pp.1446-1454.

Riloff E., S. Patwardhan and J. Wiebe. 2006. Feature Subsumption for Opinion Analysis. In *Proceedings of EMNLP-06*, pp.440-448.

Turney P. 2002. Thumbs up or Thumbs down? Semantic Orientation Applied to Unsupervised Classification of comments. In *Proceedings of ACL-02*, pp.417-424.

Vilalta R. and Y. Drissi. 2002. A Perspective View and Survey of Meta-learning. *Artificial Intelligence Review*, 18(2): 77–95.

Volkova S., W. Dolan and T. Wilson. 2012. CLex: A Lexicon for Exploring Color, Concept and Emotion Associations in Language. In *Proceedings of EACL-12*, pp.306-314.

Wilson T., J. Wiebe, and P. Hoffmann. 2009. Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis. *Computational Linguistics*, vol.35(3), pp.399-433.

Xu G., X. Meng and H. Wang. 2010. Build Chinese Emotion Lexicons Using A Graph-based Algorithm and Multiple Resources. In *Proceeding of COLING-10*, pp.1209-1217.