

# Instance Level Transfer Learning for Cross Lingual Opinion Analysis

Ruifeng Xu, Jun Xu and Xiaolong Wang

Key Laboratory of Network Oriented Intelligent Computation

Department of Computer Science and Technology

Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China

{xurui Feng, xujun}@hitsz.edu.cn, wangxl@insun.hit.edu.cn

## Abstract

This paper presents two instance-level transfer learning based algorithms for cross lingual opinion analysis by transferring useful translated opinion examples from other languages as the supplementary training data for improving the opinion classifier in target language. Starting from the union of small training data in target language and large translated examples in other languages, the Transfer AdaBoost algorithm is applied to iteratively reduce the influence of low quality translated examples. Alternatively, starting only from the training data in target language, the Transfer Self-training algorithm is designed to iteratively select high quality translated examples to enrich the training data set. These two algorithms are applied to sentence- and document-level cross lingual opinion analysis tasks, respectively. The evaluations show that these algorithms effectively improve the opinion analysis by exploiting small target language training data and large cross lingual training data.

## 1 Introduction

In recent years, with the popularity of Web 2.0, massive amount of personal opinions including comments, reviews and recommendations in different languages have been shared on the Internet. Accordingly, automated opinion analysis has attracted growing attentions. Opinion analysis, also known as sentiment analysis, sentiment classification, and opinion mining, aims to identify opinions in text and classify their sentiment polarity (Pang and Lee, 2008).

Many sentiment resources such as sentiment lexicons (e.g., SentiWordNet (Esuli and Sebastiani, 2006)) and opinion corpora (e.g., MPQA (Blitzer et al., 2007)) have been developed on different languages in which most of them are for English. The lack of reliably sentiment resources is one of the core issues in opinion analysis for other languages. Meanwhile, the manually annotation is costly, thus the amount of available annotated opinion corpora are still insufficient for supporting supervised learning, even for English. These facts motivate to “borrow” the opinion resources in one language (source language, SL) to another language (target language, TL) for improving the opinion analysis on the target language.

Cross lingual opinion analysis (CLOA) techniques are investigated to improve opinion analysis in TL through leveraging the opinion-related resources, such as dictionaries and annotated corpus in SL. Some CLOA works used bilingual dictionaries (Mihalcea et al., 2007), or aligned corpus (Kim and Hovy, 2006) to align the expressions between source and target languages. These works are puzzled by the limited aligned opinion resources. Alternatively, some works used machine translation system to do the opinion expression alignment. Banea et al. (2008) proposed several approaches for cross lingual subjectivity analysis by directly applying the translations of opinion corpus in source language to train the opinion classifier on target language. Wan (2009) combined the annotated English reviews, unannotated Chinese reviews and their translations to co-train two separate classifiers for each language, respectively.

These works directly used all of the translation of annotated corpus in source language as the training data for target language without considering the following two problems: (1) the machine translation errors propagate to following CLOA procedure; (2) The annotated corpora from different languages are collected from different domains and different writing styles which lead the training and testing data having different feature spaces and distributions. Therefore, the performances of these supervised learning algorithms are affected.

To address these problems, we propose two instance level transfer learning based algorithms to estimate the confidence of translated SL examples and to transfer the promising ones as the supplementary TL training data. We firstly apply Transfer AdaBoost (TrAdaBoost) (Dai et al., 2007) to improve the overall performance with the union of target and translated source language training corpus. A boosting-like strategy is used to down-weight the wrongly classified translated examples during iterative training procedure. This method aims to reduce the negative affection of low quality translated examples. Secondly, we propose a new Transfer Self-training algorithm (TrStr). This algorithm trains the classifier by using only the target language training data at the beginning. By automatically labeling and selecting the translated examples which is correct classified with higher confidence, the classifier is iteratively trained by appending new selected training examples. The training procedure is terminated until no new promising examples can be selected. Different from TrAdaBoost, TrStr aims to select high quality translated examples for classifier training. These algorithms are evaluated on sentence- and document-level CLOA tasks, respectively. The evaluations on simplified Chinese (SC) opinion analysis by using small SC training data and large traditional Chinese (TC) and English (EN) training data, respectively, show that the proposed transfer learning based algorithms effectively improve the CLOA. Noted that, these algorithms are applicable to different language pairs.

The rest of this paper is organized as follows. Section 2 describes the transfer learning based approaches for opinion analysis. Evaluations and discussions are presented in Section 3. Finally,

Section 4 gives the conclusions and future work.

## 2 CLOA via Transfer Learning

Given a large translated SL opinion training data, the objective of this study is to transfer more high quality training examples for improving the TL opinion analysis rather than use the whole translated training data. Here, we propose to investigate the instance level transfer learning based approaches.

In the case of transfer learning, the set of translated training SL examples is denoted by  $T_s = \{(x_i, y_i)\}_{i=1}^n$ , and the TL training data is denoted by  $T_t = \{(x_i, y_i)\}_{i=n+1}^{n+m}$ , while the size of  $T_t$  is much smaller than that of  $T_s$ , i.e.,  $|m| \ll |n|$ . The idea of transfer learning is to use  $T_t$  as the indicator to estimate the quality of translated examples. By appending selected high quality translated examples as supplement training data, the performance of opinion analysis on TL is expected to be enhanced.

### 2.1 The TrAdaBoost Approach

TrAdaBoost is an extension of the AdaBoost algorithm (Freund and Schapir, 1996). It uses boosting technique to adjust the sample weight automatically (Dai et al., 2007). TrAdaBoost joins both the source and target language training data during learning phase with different re-weighting strategy. The base classifier is trained on the union of the weighted source and target examples, while the training error rate is measured on the TL training data only. In each iteration, for a SL training example, if it is wrongly classified by prior base classifier, it tends to be a useless examples or conflict with the TL training data. Thus, the corresponding weight will be reduced to decrease its negative impact. On the contrary, if a TL training example is wrongly classified, the corresponding weight will be increased to boost it. The ensemble classifier is obtained after several iterations.

In this study, we apply TrAdaBoost algorithm with small revision to fit the CLOA task, as described in **Algorithm 1**. Noted that, our revised algorithm can handle multi-category problem which is different with original TrAdaBoost algorithm for binary classification problem only. More details and theoretical analysis of TrAdaBoost are given in Dai et al.'s work (Dai et al., 2007).

---

**Algorithm 1** CLOA with TrAdaBoost.

---

**Input:**  $T_s$ , translated opinion training data in SL,  $n = |T_s|$ ;  $T_t$ , training data in TL,  $m = |T_t|$ ;  $L$ , base classifier;  $K$ , iteration number.

- 1: Initialize the distribution of training samples:  
 $D_1(i) = 1/(n + m)$ .
- 2: **for** each  $k \in [1, K]$  **do**
- 3: Get a hypothesis  $h_k$  by training  $L$  with the combined training set  $T_s \cup T_t$  using distribution  $D_k$ :  $h_k = L(T_s \cup T_t, D_k)$ .
- 4: Calculate the training error of  $h_k$  on  $T_t$ :  
 $\epsilon_t = \sum_{i=n+1}^{n+m} \frac{D_k(i) \cdot I[h_k(x_i) \neq y_i]}{\sum_{i=n+1}^{n+m} D_k(i)}$ .
- 5: **if**  $\epsilon_t = 0$  or  $\epsilon_k \geq 1/2$  **then**
- 6:  $K = k - 1$ , break.
- 7: **end if**
- 8: Set  $\beta_k = \epsilon_k / (1 - \epsilon_k)$ ,  $\beta = 1 / (1 + \sqrt{\frac{2 \ln n}{K}})$ .
- 9: **if**  $h_k(x_i) \neq y_i$  **then**
- 10: Update the distribution:  
$$D_{k+1}(i) = \begin{cases} \frac{D_k(i)\beta}{Z_k} & 1 \leq i \leq n \\ \frac{D_k(i)/\beta_k}{Z_k} & n+1 \leq i \leq n+m \end{cases}, \text{ where}$$
  
 $Z_k$  is a normalization constant and  $\sum_{i=1}^{n+m} D_{k+1}(i) = 1$ .
- 11: **end if**
- 12: **end for**

**Output:**  $\arg \max_y \sum_{[K/2]}^K I[h_k(x) = y] \log(1/\beta_k)$   
/\*  $I[\cdot]$  is an indicator function, which equals 1 if the inner expression is true and 0 otherwise. \*/

---

## 2.2 The Transfer Self-training Approach

Different from TrAdaBoost which focuses on the filtering of low quality translated examples, we propose a new Transfer Self-training algorithm (TrStr) to iteratively train the classifier through transferring high quality translated SL training data to enrich the TL training data. The flow of this algorithm is given in **Algorithm 2**.

The algorithm starts from training a classifier on  $T_t$ . This classifier is then applied to  $T_s$ , the translated SL training data. For each category in  $T_s$  (subjective/objective or positive/negative in our different experiments), top  $p$  correctly classified translated examples are selected. These translated examples are regarded as high quality ones and thus they are appended in the TL training data. Next, the classifier is re-trained on the enriched training data. The updated classifier is applied to SL examples again to select more high quality examples. Such

---

**Algorithm 2** CLOA with Transfer Self-training.

---

**Input:**  $T_s$ , translated opinion training data in SL,  $n = |T_s|$ ;  $T_t$ , training data in TL,  $m = |T_t|$ ;  $L$ , base classifier;  $K$ , iteration number.

- 1:  $T_0 = T_t$ ,  $k = 1$ .
- 2: Get a hypothesis  $h_k$  by training a base classifier  $L$  with the training set  $T_{k-1}$ .
- 3: **for** each instance  $(x_i, y_i) \in T_s$  **do**
- 4: Use  $h_k$  to label  $(x_i, y_i)$ .
- 5: **if**  $h_t(x_i) = y_i$  **then**
- 6: Add  $(x_i, y_i)$  to  $T'$
- 7: **end if**
- 8: **end for**
- 9: Choose  $p$  instances per class with top confidence from  $T'$  and denote the set as  $T_p$ .
- 10:  $T_k = T_{k-1} \cup T_p$ ,  $T_s = T_s - T_p$ .
- 11:  $k = k + 1$ .
- 12: Iterate  $K$  times over steps 2 to 11 or repeat until  $T_p = \emptyset$ .

**Output:** Final classifier by using the enriched training set  $T_k$ .

---

procedure terminates until the increments are less than a specified threshold or the maximum number of iterations is exceeded. The final classifier is obtained by training on the union of target data and selected high quality translated SL training data.

## 3 Evaluation and Discussion

The proposed approaches are evaluated on sentence- and document-level opinion analysis tasks in the bi-lingual case, respectively. In our experiments, the TL is simplified Chinese (SC) and the SL for the two experiments are English (EN)/traditional Chinese (TC) and EN, respectively.

### 3.1 Experimental Setup

#### 3.1.1 Datasets

In the sentence-level opinionated sentence recognition experiment, the dataset is from the NTCIR-7 Multilingual Opinion Analysis Tasks (MOAT) (Seki et al., 2008) corpora. The information of this dataset is given in Table 1. Two experiments are performed. The first one is denoted by  $SenOR : TC \rightarrow SC$ , which uses  $TC_s$  as source language training dataset, while the second one

is  $SenOR: EN \rightarrow SC$ , which uses  $EN_s^1$ .  $SC_s$  is shrunk to different scale as the target language training corpus by random. The opinion analysis results are evaluated with simplified Chinese testing dataset  $SC_t$  under lenient and strict evaluation standard<sup>2</sup>, respectively, as described in (Seki et al., 2008).

| Note   | Lang. | Data     | Total | subjective/objective |          |
|--------|-------|----------|-------|----------------------|----------|
|        |       |          |       | Lenient              | Strict   |
| $SC_s$ | SC    | Training | 424   | 130/294              | \        |
| $SC_t$ |       | Test     | 4877  | 1869/3008            | 898/2022 |
| $TC_s$ | TC    | Training | 1365  | 740/625              | \        |
| $EN_s$ | EN    | Training | 1694  | 648/1046             | \        |

Table 1: The NTCIR-7 MOAT Corpora(unit:sentence).

In the document-level review polarity classification experiment, we used the dataset adopted in (Wan, 2009). Its English subset is collected by Blitzer et al. (2007), which contains a collection of 8,000 product reviews about four types of products: books, DVDs, electronics and kitchen appliances. For each type of products, there are 1,000 positive reviews and 1,000 negative ones, respectively. The Chinese subset has 451 positive reviews and 435 negative reviews of electronics products such as mp3 players, mobile phones etc. In our experiments, the Chinese subset is further split into two parts randomly: TL training dataset and test set. The cross lingual review polarity classification task is then denoted by  $DocSC: EN \rightarrow SC$ .

In this study, Google Translate<sup>3</sup> is choose for providing machine translation results.

### 3.1.2 Base Classifier and Baseline Methods

This study focus on the approaches improving the opinion analysis by using cross lingual examples, while the classifier improving on target language is not our major target. Therefore, in the experiments, a Support Vector Machines (SVM) with linear kernel is used as the base classifier. We use the

<sup>1</sup>There are only 248 sentences in NTCIR-7 MOAT English training data set. It is too small to use for CLOA. We split some samples from the test set to build a new English dataset for training, which contains all sentences from topics: N01,N02,T01,N02,N03,N04,N05,N06 and N07.

<sup>2</sup>All sentences are annotated by 3 assessors, strict standard means all 3 assessors have the same annotation and lenient means any 2 of them have the same annotation.

<sup>3</sup><http://translate.google.com/>

open source SVM package –LIBSVM(Chang and Lin, 2001) with all default parameters. In the opinionated sentence recognition experiment, we use the presences of following linguistic features to represent each sentence example including opinion word, opinion operator, opinion indicator, the unigram and bigram of Chinese words. It is developed with the reference of (Xu et al., 2008). In the review polarity classification experiment, we use unigram, bigram of Chinese words as features which is suggested by (Wan, 2009). Here, document frequency is used for feature selection. Meanwhile, term frequency weighting is chosen for document representation.

In order to investigate the effectiveness of the two proposed transfer learning approaches, they are compared with following baseline methods: (1) NoTr(T), which applies SVM with only TL training data; (2) NoTr(S), which applies SVM classifier with only the translated SL training data; (3) NoTr(S&T), which applies SVM with the union of TL and SL training data.

### 3.1.3 Evaluation Criteria

Accuracy (Acc), precision (P), recall (R) and F-measure (F1) are used as evaluation metrics. All the performances are the average of 10 experiments.

## 3.2 Experimental Results and Discussion

Here, the number of iterations in TrAdaBoost is set to 10 in order to avoid over-discarding SL examples.

### 3.2.1 Sentence Level CLOA Results

The achieved performance of the opinionated sentence recognition task under lenient and strict evaluation are given in Table 2 respectively, in which only 1/16 target train data is used as  $T_t$ . It is shown that NoTr(T) achieves a acceptable accuracy, but the recall and F1 for “subjective” category are obviously low. For the two sub-tasks, i.e.  $SenOR: TC \rightarrow SC$  and  $SenOR: EN \rightarrow SC$  tasks, the accuracies achieved by NoTr(S&T) are always between that of NoTr(T) and NoTr(S). The reason is that some translated examples from source language may likely conflict with the target language training data. It is shown that the direct use of all of the translated training data is infeasible. It is also shown that our approaches achieve better

| Method     | Sub-task | Lenient Evaluation |            |       |       |           |       |       | Strict Evaluation |            |       |       |           |       |       |
|------------|----------|--------------------|------------|-------|-------|-----------|-------|-------|-------------------|------------|-------|-------|-----------|-------|-------|
|            |          | Acc                | subjective |       |       | objective |       |       | Acc               | subjective |       |       | objective |       |       |
|            |          |                    | P          | R     | F1    | P         | R     | F1    |                   | P          | R     | F1    | P         | R     | F1    |
| NoTr(T)    |          | .6254              | .534       | .3468 | .355  | .6824     | .7985 | .7115 | .6922             | .5259      | .3900 | .3791 | .7725     | .8264 | .7776 |
| NoTr(S)    | TC → SC  | .6059              | .4911      | .7828 | .6035 | .7861     | .4960 | .6082 | .6448             | .4576      | .8352 | .5912 | .8845     | .5603 | .6860 |
| NoTr(S&T)  |          | .6101              | .4943      | .7495 | .5957 | .7711     | .5236 | .6235 | .6531             | .4632      | .8051 | .588  | .8714     | .5856 | .7004 |
| TrAdaBoost |          | .6533              | .5335      | .7751 | .6314 | .8063     | .5777 | .6720 | .7184             | .5273      | .8473 | .6494 | .9077     | .6611 | .7643 |
| TrStr      |          | <b>.6625</b>       | .5448      | .7309 | .6238 | .7884     | .6199 | .6934 | <b>.7304</b>      | .5414      | .8182 | .6511 | .896      | .6914 | .7801 |
| NoTr(S)    | EN → SC  | .6590              | .5707      | .4446 | .4998 | .6966     | .7922 | .7413 | .7390             | .5872      | .5100 | .5459 | .7944     | .8408 | .8169 |
| NoTr(S&T)  |          | .6411              | .5292      | .5759 | .5515 | .7212     | .6817 | .7009 | .7105             | .5254      | .608  | .5637 | .8129     | .7560 | .7834 |
| TrAdaBoost |          | <b>.6723</b>       | .5988      | .4371 | .5018 | .7019     | .8184 | .7549 | <b>.7630</b>      | .6485      | .5019 | .5614 | .8002     | .8789 | .8371 |
| TrStr      |          | .6686              | .5691      | .5746 | .5678 | .7360     | .7271 | .7292 | .7484             | .589       | .6276 | .6026 | .8315     | .8021 | .8147 |

Table 2: The Performance of Opinionated Sentence Recognition Task.

performance on both tasks while few TL training data is used. In which, TrStr performs the best on  $SenOR:TC \rightarrow SC$  task while TrAdaBoost outperforms other methods on  $SenOR:EN \rightarrow SC$  task. The proposed transfer learning approaches enhanced the accuracies achieved by NoTr(S&T) for 4.2-8.6% under lenient evaluation and 5.3-11.8% under strict evaluation, respectively.

### 3.2.2 Document Level CLOA Results

| Method     | Acc          | positive |       |       | negative |       |       |
|------------|--------------|----------|-------|-------|----------|-------|-------|
|            |              | P        | R     | F1    | P        | R     | F1    |
| NoTr(T)    | .7542        | .7447    | .8272 | .7747 | .8001    | .6799 | .7235 |
| NoTr(S)    | .7122        | .6788    | .8248 | .7447 | .7663    | .5954 | .6701 |
| NoTr(S&T)  | .7531        | .714     | .8613 | .7801 | .8187    | .6415 | .7179 |
| TrAdaBoost | .7704        | .8423    | .6594 | .7376 | .7285    | .8781 | .7954 |
| TrStr      | <b>.7998</b> | .8411    | .7338 | .7818 | .7727    | .8638 | .8144 |

Table 3: The Results of Chinese Review Polarity Classification Task (Features:Unigrams; m=20).

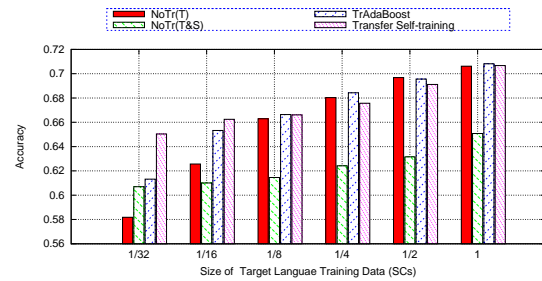
| Method     | Acc          | positive |       |       | negative |       |       |
|------------|--------------|----------|-------|-------|----------|-------|-------|
|            |              | P        | R     | F1    | P        | R     | F1    |
| NoTr(T)    | .7518        | .7399    | .8294 | .7741 | .7983    | .6726 | .7185 |
| NoTr(S)    | .7415        | .7143    | .8204 | .7637 | .7799    | .6598 | .7148 |
| NoTr(S&T)  | .7840        | .7507    | .8674 | .8035 | .8385    | .6982 | .7592 |
| TrAdaBoost | .7984        | .8416    | .7297 | .7792 | .7707    | .8652 | .8138 |
| TrStr      | <b>.8022</b> | .8423    | .7393 | .7843 | .7778    | .8634 | .8164 |

Table 4: The Results of Chinese Review Polarity Classification Task (Features:Unigrams+Bigrams; m=20).

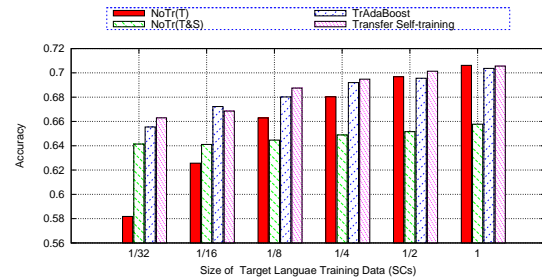
Table 3 and Table 4 give the achieved results of different methods on the task  $DocSC:EN \rightarrow SC$  by using 20 Chinese annotated reviews as  $T_t$ . It is shown that transfer learning approaches outperform

other methods, in which TrStr performs better than TrAdaBoost when unigram-bigram features are used. Compared to NoTr(T&S), the accuracies are increased about 1.8-6.2% relatively. Overall, the transfer learning approaches are shown are beneficial to TL polarity classification.

### 3.2.3 Influences of Target Training Corpus Size



(a)  $SenOR:TC \rightarrow SC$



(b)  $SenOR:EN \rightarrow SC$

Figure 1: Performances with Different Size of  $SC_s$  on Opinionated Sentence Recognition Task under Lenient Evaluation

In order to estimate the influence of different size of TL training data, we conduct a set of experiments on both tasks. Fig 1 and Fig 2 show the influence

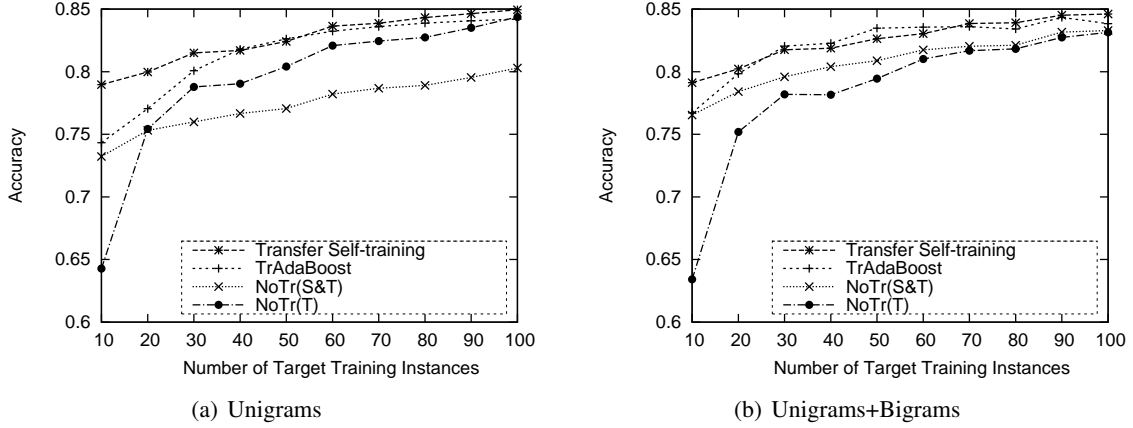


Figure 3: Performances with Different Number of TL Training Instances on Task of *DocSC: EN → SC*

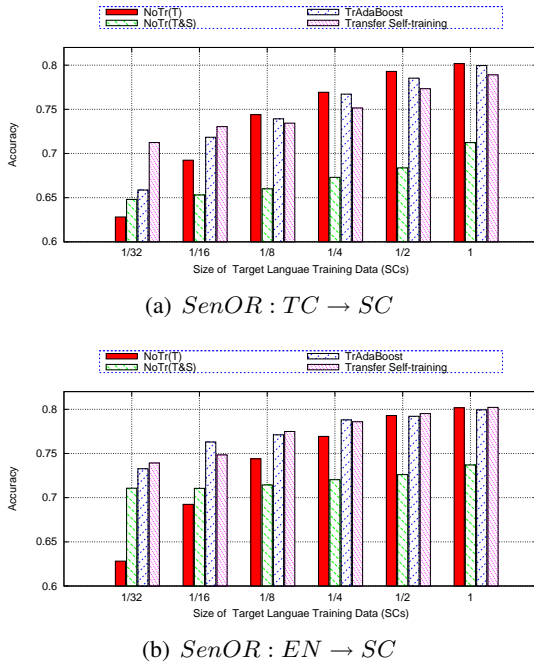


Figure 2: Performances with Different Size of  $SC_s$  on Opinionated Sentence Recognition Task under Strict Evaluation

on the opinionated sentence recognition task under lenient and strict evaluation respectively. Fig 3 shows the influence on task *DocSC: EN → SC*. Fig 3(a) shows the results use unigram features and Fig 3(b) uses both unigrams and bigrams. It is observed that TrAdaBoost and TrStr achieve better performances than the baseline NoTr(S&T) in most cases. More specifically, TrStr performs the best when few TL training data is used. When more TL

training data is used, the performance improvements by transfer learning approaches become small. The reason is that less target training data is helpful to transfer useful knowledge in translated examples. If too much TL training data is used, the weights of SL instances may decrease exponentially after several iterations, and thus more source training data is not obviously helpful.

#### 4 Conclusions and Future Work

To address the problems in CLOA caused by inaccurate translations and different domain/category distributions between training data in different languages, two transfer learning based algorithms are investigated to transfer promising translated SL training data for improving the TL opinion analysis. In this study, Transfer AdaBoost and Transfer Self-Training algorithms are investigated to reduce the influences of low quality translated examples and to select high quality translated examples, respectively. The evaluations on sentence- and document-level opinion analysis tasks show that the proposed algorithms improve opinion analysis by using the union of few TL training data and selected cross lingual training data.

One of our future directions is to develop other transfer learning algorithms for CLOA task. Another future direction is to employ other moderate weighting scheme on source training dataset to reduce the over-discarding of training examples from source language.

## References

- Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1-135.
- Andrea Esuli and Fabrizio Sebastiani. 2006. SENTIWORDNET: A publicly available lexical resource for opinion mining. *Proceedings of the 5th International Conference on Language Resources and Evaluation*, 417-422.
- Janyce Wiebe, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 39(2-3):165-210.
- Rada Mihalcea, Carmen Banea, and Janyce Wiebe. 2007. Learning multilingual subjective language via cross-lingual projections. *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, Prague, Czech Republic.
- Soo-Min Kim and Eduard Hovy. 2006. Identifying and analyzing judgment opinions. *Proceedings of HLT/NAACL-2006*, 200-207.
- Carmen Banea, Rada Mihalcea, Janyce Wiebe and Samer Hassan. 2008. Multilingual subjectivity analysis using machine translation. *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, Honolulu, Hawaii, 127-135.
- Xiaojun Wan. 2009. Co-training for cross-lingual sentiment classification. *Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP*, Suntec, Singapore, 235-243.
- Wenyuan Dai ,Qiang Yang, GuiRong Xue and Yong Yu. 2007. Boosting for transfer learning. *Proceedings of the 24th International Conference on Machine Learning*, 193-200.
- John Blitzer, Mark Dredze, and Fernando Pereira. 2007. Biographies, bollywood, boom-boxes and blenders: domain adaptation for sentiment classification. *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, 440-447.
- Xiaojun Wan. 2008. Using bilingual knowledge and ensemble techniques for unsupervised chinese sentiment analysis. *Proceedings of EMNLP 2008*,553-561.
- Yoav Freund and Robert E. Schapire. 1996. Experiments with a new boosting algorithm. *Proceedings of the 13th International Conference on Machine Learning*, 148-156.
- Yohei Seki, David K. Evans, Lun-Wei Ku, Le Sun, Hsin-Hsi Chen, and Noriko Kand. 2008. Overview of multilingual opinion analysis task at NTCIR-7. *Proceeding of NTCIR-7*, NII, Tokyo, 185-203.
- Ruifeng Xu, Kam-Fai Wong, Qin Lu, and Yunqing Xia. 2008. Learning Multilinguistic Knowledge for Opinion Analysis. D. S. Huang et al., editors:*Proceedings of ICIC 2008*, volume 5226 of LNCS, 993-1000, Springer-Verlag.
- Chih-Chung Chang and Chih-Jen Lin. 2001. LIBSVM: a library for support vector machines. *Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>*.