

Singapore Management University

## Institutional Knowledge at Singapore Management University

---

Research Collection School Of Information  
Systems

School of Information Systems

---

4-2018

### Personalized microblog sentiment classification via adversarial cross-lingual learning

Weichao WANG

Shi FENG

Wei GAO

Singapore Management University, weigao@smu.edu.sg

Daling WANG

Yifei ZHANG

Follow this and additional works at: [https://ink.library.smu.edu.sg/sis\\_research](https://ink.library.smu.edu.sg/sis_research)



Part of the [Databases and Information Systems Commons](#)

---

#### Citation

WANG, Weichao; FENG, Shi; GAO, Wei; WANG, Daling; and ZHANG, Yifei. Personalized microblog sentiment classification via adversarial cross-lingual learning. (2018). *Proceedings of 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 338-348. Research Collection School Of Information Systems.

Available at: [https://ink.library.smu.edu.sg/sis\\_research/4560](https://ink.library.smu.edu.sg/sis_research/4560)

This Conference Proceeding Article is brought to you for free and open access by the School of Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email [libIR@smu.edu.sg](mailto:libIR@smu.edu.sg).

# Personalized Microblog Sentiment Classification via Adversarial Cross-lingual Multi-task Learning

Weichao Wang<sup>1</sup>, Shi Feng<sup>1</sup>, Wei Gao<sup>2</sup>, Daling Wang<sup>1</sup>, Yifei Zhang<sup>1</sup>

<sup>1</sup>School of Computer Science and Engineering, Northeastern University, Shenyang, China  
wangwecha@gmail.com, {fengshi, wangdaling, zhangyifei}@cse.neu.edu.cn

<sup>2</sup>Victoria University of Wellington, Wellington, New Zealand  
wei.gao@vuw.ac.nz

## Abstract

Sentiment expression in microblog posts can be affected by user's personal character, opinion bias, political stance and so on. Most of existing personalized microblog sentiment classification methods suffer from the insufficiency of discriminative tweets for personalization learning. We observed that microblog users have consistent individuality and opinion bias in different languages. Based on this observation, in this paper we propose a novel user-attention-based Convolutional Neural Network (CNN) model with adversarial cross-lingual learning framework. The user attention mechanism is leveraged in CNN model to capture user's language-specific individuality from the posts. Then the attention-based CNN model is incorporated into a novel adversarial cross-lingual learning framework, in which with the help of user properties as bridge between languages, we can extract the language-specific features and language-independent features to enrich the user post representation so as to alleviate the data insufficiency problem. Results on English and Chinese microblog datasets confirm that our method outperforms state-of-the-art baseline algorithms with large margins.

## 1 Introduction

The past decade has witnessed an exploding interest in microblog sentiment classification from both academic and commercial communities due to its inherent challenges and wide applications (Liu, 2012; Jawa and Hasija, 2015). However, most existing algorithms largely ignore latent personal distinctions among microblog users (Wu et al., 2015; Sun et al., 2015; Yu et al., 2015). In fact, the sentiment expression in posts can be affected by user's language habit, personal character, opinion bias and so on, and even the same sentence can deliver quite different sentiment polarities depending on user's underlying context.

Table 1 illustrates a simple case where two users are talking about house price. In the first two example posts, we can see that the **Seller**, as a real estate agent, is excited about the growth of the price, while as a consumer, the **Buyer** feels sad for the situation. Without knowing the users' background, it would be difficult for the traditional methods to detect the sentiment polarities of posts 1 and 2.

In recent years, personalized sentiment classification considering user's individuality has emerged as one of the prevailing research topics in the field. Song et al. (2015b, 2016) utilized a latent factor model to capture personal sentiment variations of microblog users. Zhao et al. (2017) exploited the social relation information to facilitate user representation learning. In rating prediction research, Tang et al. (2015) embedded user and product level information into a neural network model. Chen et al. (2016a) further improved the model by incorporating global user and product information as attentions into a hierarchical LSTM model. These studies have achieved promising results on personalized sentiment classification. However, most of them suffer from the insufficiency of discriminative posts for personalization learning.

We observed that the microblog users have consistent individuality and opinion bias in different languages on different platforms. Posts 3 and 4 in Table 1 are posted by the same couple of **Seller** and **Buyer** in Chinese on Sina Weibo<sup>1</sup>. Both of them expressed their own stance or opinion across the two platforms as consistently as what each has said in English on Twitter, indicating the stable underlying individuality of each user. Jointly analyzing microblog posts in different languages can help us enrich the user features and ex-

<sup>1</sup>[www.weibo.com](http://www.weibo.com)

ID	Source	Post	Label
1	Twitter	<b>Seller:</b> The house prices are getting higher and higher!	(positive)
2	Twitter	<b>Buyer:</b> The house prices are getting higher and higher!	(negative)
3	Weibo	<b>Seller:</b> 房价一直在提高, 我们会有不错的效益。 ( <i>The house prices have been rising and we will have good benefits.</i> )	(positive)
4	Weibo	<b>Buyer:</b> 过高的房价已经超出了我的承受范围。 ( <i>High house prices have exceeded my tolerance.</i> )	(negative)

Table 1: A toy example of microblogs of the same users in English and Chinese. Note that posts 3 and 4 are brand new and not the translation of posts 1 and 2.

tract more precise personalized view and opinion bias. Furthermore, the microblog posts on Twitter and Weibo share many similarities in social and textual styles regardless of different languages used. By cross-lingual learning from the same (or similar) users’ posts, we can extract the common language-independent features to enhance the representations for both monolingual and cross-lingual sentiment classification. We hypothesize that general cross-lingual sentiment words such as adjectives have similar effect on the sentiment of sentences in different languages, which can be considered generally independent of language. It is thus meaningful to leverage the individuality features to bridge the language gap and exploit the global language-independent sentiment knowledge learned to help alleviate data insufficiency for personalized sentiment classification.

Convolutional neural networks (CNN) are effective in short text analysis thanks to the  $n$ -grams of text that contain rich and accurate information for capturing user’s attention. Besides, adversarial multi-task learning provides us with a feasible method to bridge the semantics of posts of different languages by utilizing the correlation between the two related language-specific tasks. In this paper, we propose a user-attention-based CNN model with adversarial multi-task learning (dubbed as ACNN-AMT). Firstly, we incorporate user attention into CNN model to capture user’s language-specific individuality from the posts. Secondly, we propose an adversarial multi-task learning framework to co-train cross-lingual sentiment features with the help of user’s consistent individuality. The main contributions of our paper are three-fold:

- We propose a novel user attention mechanism over  $n$ -gram semantics of short posts to capture the language-specific individuality and opinion bias of users for personalized senti-

ment classification.

- We incorporate the user-attention-based CNN model into a novel adversarial cross-lingual multi-task learning framework. With the help of consistency of user individuality as bridge across two languages, we learn language-specific features and language-independent features to enrich the user post representation, and alleviate the problem of insufficiency of personal data.
- We conduct comprehensive experiments on real-world Sina Weibo datasets and Twitter datasets to confirm the effectiveness of our proposed model. The experimental results demonstrate that our model consistently outperforms strong baseline methods by a large margin. Our datasets have been made publicly available<sup>2</sup>.

## 2 Related Work

Recently, personalized sentiment classification has attracted lots of attention in research community. Song et al. (2015b) proposed a personalized sentiment classification method based on latent factor model to capture personal sentiment variations. Tang et al. (2015) was first to incorporate user and product information into a neural network model for personalized rating prediction of products. Wu and Huang (2016) proposed a personalized microblog sentiment classification framework via multi-task learning to capture the personal properties and take advantage of common sentiment knowledge shared among all users. Chen et al. (2016a) adopted a hierarchical neural network to incorporate global user and product information into the sentiment model via attention mechanism. Zhao et al. (2017) exploited both the

<sup>2</sup><https://github.com/githubfordata/data>

semantics of posts and the social relations among users for microblog sentiment classification with recurrent random-walk network model. Although these methods have achieved reasonable results, they are limited by insufficient personalized data that the learning of personal sentiment features and user representations has to rely on.

Generative adversarial network (Goodfellow et al., 2014) has been successfully applied to Natural Language Processing, especially domain and language adaptation, including measuring and matching different feature space distribution between source and target. Ganin and Lempitsky (2015) presented an adversarial approach to domain adaptation for transferring knowledge from source domain to target domains. Chen et al. (2016b) proposed an adversarial deep averaging network to transfer sentiment knowledge from labeled source language to low-resource target language where only unlabeled data exists. Liu et al. (2017) tried an adversarial multi-task learning framework for text classification, intending to refrain the shared and private latent feature spaces from interfering each other. These methods only took multi-domain or multi-language setting into account, aiming to address the problem caused by low-resource languages or specific domains, but they ignored the important factors of users across different domains or languages, which is a very important factor in sentiment analysis.

Unlike the previous studies, we incorporate cross-lingual sentiment consistency and personalized user sentiment model from the viewpoint of multi-task learning for alleviating the insufficiency of personal data and improving the performance of personalization in sentiment analysis.

### 3 Attention-based Personalization Model

In CNN, we can use attention mechanism to predict the importance probability of each component with respect to attention query (Bahdanau et al., 2014). Given a post, the user usually focuses on its specific portions depending on one’s individuality or interest. The user attention can be incorporated into CNN model to learn appropriate user-specific features and capture crucial components that the user cares about.

Given a data example  $(x, u, y)$  in a monolingual dataset, the post content  $x$  is denoted as:

$$x = w_1 \oplus w_2 \oplus \dots \oplus w_i \oplus \dots \oplus w_n \quad (1)$$

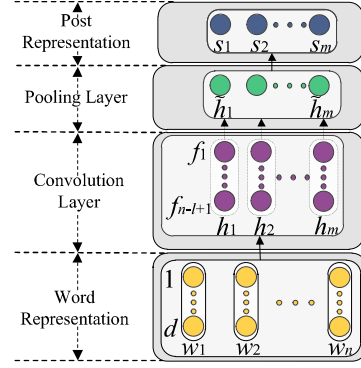


Figure 1: Convolutional Neural Network for post representation.

where  $n$  is number of words contained in  $x$ ,  $\oplus$  is the concatenation operator, each word  $w_i$  is mapped to its  $d$ -dimensional embedding representation  $w_i \in \mathbb{R}^d$ , the user of the post is represented by a continuous vector of  $o$  dimensions  $u \in \mathbb{R}^o$  which is initialized as zero-valued vector, and  $y$  is the sentiment label of  $x$ . Given all training examples, our goal is to learn appropriate text and user representations for classifying sentiment class of different users’ new posts.

#### 3.1 CNN-based Text Representation

We use CNN (Kim, 2014; Kalchbrenner et al., 2014) for learning the semantic representation  $s$  of input post  $x$ , which is shown in Figure 1. The semantic representations of  $n$ -grams with sliding window are mapped to fixed and lower dimension via the convolutional filters. Let  $l$  be the width of convolutional filter,  $W$  and  $b$  be the parameters and bias of linear mapping, which is shared among all layers in a filter. Each sliding window with fixed-length window size of  $l$  starting from the  $k$ -th word is denoted as  $\Omega_k = w_k \oplus w_{k+1} \oplus \dots \oplus w_{k+l-1}$ , where  $\Omega_k \in \mathbb{R}^{d \times l}$ . The convolutional layer output  $f_k \in \mathbb{R}^m$  of the window  $\Omega_k$  is calculated by applying the linear mapping below, following a convolution operation:

$$f_k = W \cdot \Omega_k + b \quad (2)$$

where  $W \in \mathbb{R}^{m \times d}$ ,  $b \in \mathbb{R}^m$ , and  $m$  is the output length of each linear mapping. With the convolution operation,  $n - l + 1$  features in total, i.e., one for each window, are obtained and denoted as  $\mathbf{f} = [f_1, f_2, \dots, f_{n-l+1}]^T$ . For each column vector  $h_j$  for  $j = 1, \dots, m$  in  $\mathbf{f}$ , a scalar  $\tilde{h}_j$  is calculated as the representation of  $h_j$ . By concatenating all scalars, the output vector of a post representation

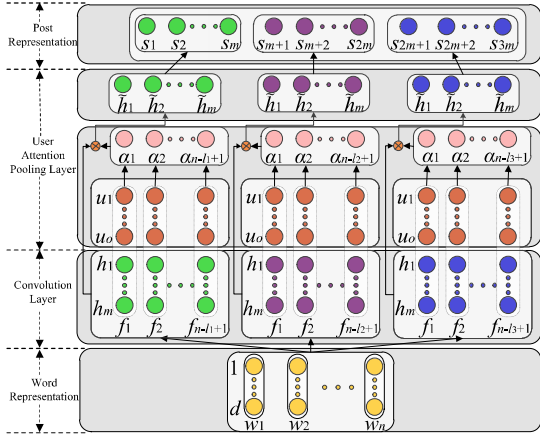


Figure 2: Our user-attention-based CNN model (ACNN) for post representation. The convolution windows (in convolution layer) in green, purple and blue color correspond to three sizes of  $n$ -grams used, i.e.,  $l_1$ ,  $l_2$ , and  $l_3$ , respectively.

$s = [s_1, s_2, \dots, s_m]$  with fixed length of  $m$  is generated. Note that  $s$  can be generated from  $\tilde{h}$  with the similar linear mapping as Eq. 2, but for simplicity we treat it the same as  $\tilde{h}$  in our experiment.

### 3.2 Incorporating User Attention

The model incorporating user attention mechanism is shown in Figure 2.

The users information are distinguished by user id, and each user  $u$  is embedded as vector  $u \in \mathbb{R}^o$  for capturing their opinion bias and individuality. All the users embedding will be regarded as attention query parts in user attention mechanism, and the model gives all input windows different attention weights according to the individuality of post users for semantic representation. We treat user embeddings as model parameters, which can be updated with the optimization algorithm in training process (see Section 4).

Formally, the final post representation  $s = \tilde{h}$  is a weighted sum of all sliding windows of convolution layer:

$$\tilde{h} = \sum_{k=1}^{n-l+1} \alpha_k f_k \quad (3)$$

where  $\alpha_k$  measures the importance of the  $k$ -th  $n$ -gram (i.e., window) for current user. The attention weight  $\alpha_k$  of each window's hidden state can be defined as:

$$\alpha_k = \frac{\exp(\sigma(f_k, u))}{\sum_{j=1}^{n-l+1} \exp(\sigma(f_j, u))} \quad (4)$$

where  $\sigma$  is a scoring function which scores the importance of each  $n$ -gram by taking into account both post and user representations.

For each input window, the attention score is calculated for getting the attentive degree of input to current user attention query by combining them together linearly. The scoring function  $\sigma$  is defined as:

$$\sigma(f_k, u) = v \cdot \tanh(W_f \cdot f_k + W_u \cdot u + b) \quad (5)$$

where  $W_f \in \mathbb{R}^{m \times m}$  is a weight matrix for hidden states of input windows of a post,  $W_u \in \mathbb{R}^{m \times o}$  is weight matrix for the user of the current post, and  $v \in \mathbb{R}^m$  is the weight vector, which is randomly initialized, for getting a score number for the function.  $v$  is needed since a matrix  $\mathbb{R}^{m \times 1}$  is obtained by  $\tanh(\cdot)$ , and for getting a scalar, we need a matrix  $\mathbb{R}^{1 \times m}$  for matrix multiplication.

To make our model more accurate, we apply three kinds of convolutional filters to capture the semantics of  $n$ -grams corresponding to three levels of granularities, i.e., we let  $l \in \{l_1, l_2, l_3\}$  corresponding to unigram, bi-gram and tri-gram, respectively, and we concatenate the representations  $\tilde{h}$  of these  $n$ -grams of these three granularities as the representation of each post. Similar as (Tang et al., 2015), for modeling user-sentiment consistencies, we also concatenate the representation of users and that of posts as the final semantic representation. Note that for the current post, the user embedding is the same among three different granularities, but the model parameters  $W_f$  and  $W_u$  among the three granularities are different because the convolutional representation of post  $f_k$  is dependent on specific granularity of  $n$ -grams used.

## 4 Adversarial Learning for Personalized Sentiment Classification

Multi-task learning can utilize the correlation among different tasks to learn related knowledge for better performance (Evgeniou and Pontil, 2004). The posts of each user are often insufficient for personalized sentiment analysis with a monolingual dataset. By exploiting posts of both Chinese and English in multi-task learning, we can enrich the original monolingual posts of each user to alleviate the data insufficiency problem. Inspired by the adversarial network that can build the mapping between generative distribution and target distribution (Goodfellow et al., 2014),



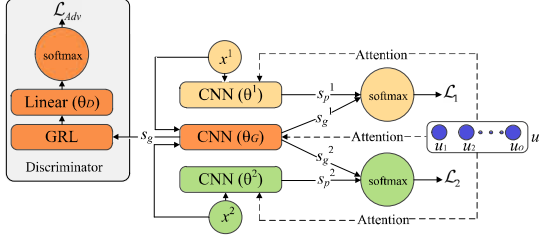


Figure 3: The ACNN-AMT model for sentiment classification. The left hand side is the language discriminator of the adversarial network; the CNNs in the middle are the feature generators of language-specific sentiment classification tasks; and the right hand side are the decision functions of the corresponding sentiment classifiers. Both sentiment classification tasks share the same user representation where the bridge is established via user’s consistent individuality across languages. The adversarial learning is realized by gradient reversal layer (GRL).

we conjecture that the sentiment knowledge between two different languages can be mapped to the same feature space to further improve the effect of multi-task learning with a feature generator and a language category discriminator. Formally, given a data sample  $(x^t, u^t, y^t)$  from the posts of the  $t$ -th language (i.e., task), we aim to predict the sentiment class of the posts with our adversarial cross-lingual multi-task learning method.

#### 4.1 Adversarial Multi-task Learning

Intuitively, a user may post on different microblogging platforms in different languages, but the individuality and opinion bias embedded in the posts of different languages are supposed to be consistent. By regarding user’s individuality as bridge between languages, we adopt multi-task learning to utilize the correlation of user’s individuality across Chinese and English platforms for learning better user representation.

As shown in Figure 3, we illustrate our user attention based CNN model with adversarial multi-task learning (ACNN-AMT). All tasks share a same collection of users embedding  $u$  via the attention mechanism as a cross-lingual bridge. Each task is a sentiment classification model, which is decomposed into two components, i.e., a global ACNN model trained with posts from all languages for extracting common sentiment features (the orange-color CNN model), a language-specific ACNN model which is trained with posts from a specific language (the yellow- or green-color CNN model) for generating language-

specific features, and a softmax classifier for predicting sentiment classification. Therefore, this multi-task model is both a sentiment classifier and meanwhile a generator for producing post representations at global and language-specific levels. Both tasks share the same global ACNN model.

For any input post  $x^t$  of the  $t$ -th task<sup>3</sup>, we can compute its global representation  $s_g^t$  and task-specific representation  $s_p^t$  with the ACNN model. The final features  $s^t$  of post  $x^t$  are concatenation of the features from the specific and global space and user embedding. Formally, we define  $s_g^t$ ,  $s_p^t$  and  $s^t$  as follows:

$$\begin{aligned} s_g^t &= \text{ACNN}(x^t, u^t, \theta_G) \\ s_p^t &= \text{ACNN}(x^t, u^t, \theta^t) \\ s^t &= s_g^t \oplus s_p^t \oplus u^t \end{aligned} \quad (6)$$

where  $\theta_G$  contain parameters of the global task-independent model (which acts as the generator in the adversarial network), and  $\theta^t$  contain the model parameters of the  $t$ -th task, and  $u^t$  is the user embedding of the  $t$ -th task which is constant across different tasks.

Because Chinese and English have different lexical and linguistic structures, the global sentiment classifier could learn little knowledge from English for Chinese sentiment classification, and vice versa. Nevertheless, if we could learn sentiment features independent of language category, the global sentiment classifier will be well trained by posts of both languages. To get such language-independent features, we introduce adversarial training into the global model. We formulate the global model as a generator and the language classification model as a discriminator using Generative Adversarial Network (GAN) (Goodfellow et al., 2014). During training, the generator is trying to generate language-independent sentiment features that would confuse the discriminator as to these features’ language category, so that the discriminator is forced to struggle to become a strong classifier for correctly predicting the language category of the generated representation, until both the generator and discriminator cannot improve any more. The language discriminative model  $D$  parameterized by  $\theta_D$  and  $\theta_G$  is defined as follows:

$$\hat{y}(\theta_D, \theta_G) = \text{softmax}(\tanh(W_D s^g(\theta_G) + b_D)) \quad (7)$$

<sup>3</sup>For simplicity, we consider a two-task case here, i.e.,  $t \in \{1, 2\}$ . But our model is generic which can easily consider more tasks.

where  $\hat{y}(\cdot)$  is the predicted language category,  $\theta_D$  contain parameters  $W_D \in \mathbb{R}^{T \times 3m}$  which are the weights ( $T$  is number of language categories) and  $b_D \in \mathbb{R}^{T \times 3m}$  which are the bias terms, and  $s^g(\theta_G) \in \mathbb{R}^{3m}$  is the language-independent sentiment features extracted by global ACNN model parameterized by  $\theta_G$ .

Finally, we get the adversarial loss by cross-entropy error between gold language distribution and our predicted language distribution:

$$\mathcal{L}_{Adv} = \sum_{t=1}^T \sum_{i=1}^{N^t} y_i^t \log(\hat{y}_i^t) \quad (8)$$

where  $y_i^t$  denotes the ground-truth language label indicating that the  $i$ -th post comes from the  $t$ -th task (or language),  $\hat{y}_i^t$  is the predicted language of the post, and  $N^t$  is the number of posts in the  $t$ -th task. We use the gradient reversal layer (GRL) (Ganin and Lempitsky, 2015) to perform the adversarial learning. The objective of GRL is to optimize  $\theta_G$  and  $\theta_D$  via the min-max steps as  $\min_{\theta_G} \max_{\theta_D} \mathcal{L}_{Adv}$ . After the adversarial training, we can obtain stronger language category discriminator and language-independent sentiment features.

## 4.2 Sentiment Classification

For the  $t$ -th task, the predicted post sentiment distribution is defined as follows:

$$\hat{z}^t = \text{softmax}(\tanh(W^t s^t + b^t)) \quad (9)$$

where  $W^t \in \mathbb{R}^{C \times (6m+o)}$  are the weights and  $b^t \in \mathbb{R}^{(6m+o)}$  are the biases. For each sentiment category, cross-entropy error between gold sentiment distribution and our predicted sentiment distribution is defined as sentiment loss:

$$\mathcal{L}^t = - \sum_{c=1}^C \sum_{i=1}^{N_c^t} z_{c,i}^t \log(\hat{z}_{c,i}^t) \quad (10)$$

where  $z_{c,i}^t$  denotes the ground-truth label indicating sentiment of the  $i$ -th post of the  $t$ -th task,  $\hat{z}_{c,i}^t$  is the predicted sentiment label of the post,  $C$  is the number of sentiment classes, and  $N_c^t$  is the number of posts of class  $c$  in the  $t$ -th task.

Similar to (Liu et al., 2017), we add a squared Frobenius norm term to the loss function:

$$\mathcal{L}_{Frobenius} = \sum_{t=1}^T \|G^{t\top} P^t\|_F^2 \quad (11)$$

where  $\|\cdot\|_F^2$  is the squared Frobenius norm,  $G^t$  and  $P^t$  are matrices whose rows are  $s_g^t$  and  $s_p^t$  of each input sentence, respectively.

The final loss function of our model can be written as:

$$\mathcal{L} = \beta \mathcal{L}^1 + \lambda \mathcal{L}^2 + \gamma \mathcal{L}_{Adv} + \delta \mathcal{L}_{Frobenius} \quad (12)$$

where  $\mathcal{L}^1$  and  $\mathcal{L}^2$  are the loss functions of sentiment classification models of the two languages, and the hyper parameters  $\beta$ ,  $\lambda$ ,  $\gamma$  and  $\delta$  are the weights of their corresponding losses.

To minimize the overall loss function  $\mathcal{L}$ , we use Adadelta (Zeiler, 2012) to update the model parameters. In particular, we employ Adadelta to update  $\theta_D$  to maximize the adversarial loss term  $\mathcal{L}_{Adv}$ , and update other parameters to minimize the overall loss function  $\mathcal{L}$ .

## 5 Experiment

### 5.1 Experimental Setting

Our experiments are conducted on two real-world microblog sentiment datasets. The first dataset was crawled using our built-in-house crawler from Sina Weibo, which contain only Chinese posts. Weibo users have their personal profile, and their homepage links to Twitter may also appear in it. We extracted the Twitter homepage links for crawling our Twitter dataset via Twitter API, and the dataset contain only English posts. We labeled the posts with sentiment polarity based on frequently used emoticons as emoticons which are given by users were commonly regarded as appropriate personal sentiment labels (Song et al., 2015a,b; Wu and Huang, 2016).

The statistics of the two datasets are summarized in Table 2 and 3, where 96 bilingual users were found appearing in both datasets. In the experiment, we used all users, i.e., 755 from Twitter, and 379 from Weibo, assuming that similar users can be embedded into similar user representation space. The bilingual users will especially bring strong effects on personalization learning because the posts of bilingual users can be better represented with consistent individuality than those of monolingual users.

For each user's posts in different languages, we split them into training, development and testing sets by user volumes under the ratio of 7:2:1 randomly, and used Stanford CoreNLP (Manning et al., 2014) for word segmentation of Chinese and English posts. For evaluation, we use precision,

Dataset	classes	users	posts	positive	negative	posts/user	vocabulary	words/post
Twitter	2	755	41,210	28,782	12,428	54.58	27,211	10.70
Weibo	2	379	40,081	25,367	14,714	105.75	25,702	15.74

Table 2: Statistics of Weibo and Twitter datasets.

Dataset	users	posts	positive	negative	posts/user	vocabulary	words/post
Twitter	96	6,942	4,879	2,063	72.31	10,181	9.84
Weibo		13,563	7,448	6,115	141.28	14,031	15.19

Table 3: Statistics of Weibo and Twitter data of bilingual users.

recall and F-score to measure the classification results with 10-fold cross validation, and take average over ten folds for obtaining the values of the three evaluation metrics.

For Chinese posts, we trained our Word2Vec model (Mikolov et al., 2013) on our crawled 30M Weibo corpus. For English posts, we used the 200d GloVe vectors (Pennington et al., 2014) as word embeddings. We set the dimensionality of user embedding to 100.

The other model parameters were initialized by randomly sampling from uniform distribution in  $[-0.1, 0.1]$ . We set the mini-batch size to 32. Besides, we employed dropout technique (Srivastava et al., 2014) and the dropout rate was set to 0.5. We tuned the hyper-parameters on the development sets, and empirically set  $\beta$  to 0.8,  $\lambda$  to 1,  $\gamma$  to 0.1, and  $\delta$  to 0.01. The width of three convolutional filter was fixed to 1, 2, and 3, corresponding to unigram, bi-gram and tri-gram, respectively.

## 5.2 Baselines

We compared our model with several strong baseline methods for microblog sentiment classification:

**SVM + Wordvec:** It averages the embeddings of all words in a post, which are regarded as features of a SVM classifier using LibSVM library (Chang and Lin, 2011).

**LSTM:** It uses traditional LSTM model for sentiment classification.

**CNN:** It uses CNN model (Kim, 2014) for sentiment classification.

**NSC + UPA:** It uses user and product attention mechanism over word-level and sentence-level semantics based on LSTM (Chen et al., 2016a). The original method focused on modeling product review texts. Because microblog posts contain users but no particular products information, we imple-

mented the method with word-level modeling and user attention for fair comparison.

**UPNN:** It uses vector space model for modeling user preferences and product qualities, which is integrated with word-level representation and post-level representation (Tang et al., 2015). Original UPNN models user preference and product characteristic. We modified it to only model user preferences as a baseline method for comparison.

**ASP-MTL:** It uses LSTM model with adversarial multi-task learning to extract the domain-invariant features for cross-domain text classification (Liu et al., 2017). We regard domain category as language category for cross-lingual sentiment classification.

## 5.3 Result Comparisons

Table 4 shows the comparison of results among different models. We separate results into two groups which are monolingual results and cross-lingual results respectively.

In the first group, the SVM with averaging word embedding performs poorly because SVM cannot capture semantic information as well as deep neural networks. The CNN model performs better than LSTM, because LSTM model is easier to overfit. The NSC with UPA yields better results than LSTM, and UPNN yields better results than CNN, because they exploit the impact of user information. Our ACNN model archives the best performance by exploiting user information on top of the semantics of  $n$ -grams.

In the second group, the ASP-MTL performs better than LSTM, because global sentiment classifier is well trained by language-independent sentiment features with adversarial training. Our ACNN-AMT model significantly outperforms all the other baseline methods, confirming the importance of both user information and global



Models	Twitter			Weibo		
	P	R	F	P	R	F
SVM+Wordvec	77.18	68.74	70.63	72.60	68.07	68.88
CNN	75.60	77.02	76.41	73.57	73.60	73.55
LSTM	74.02	75.15	74.59	71.11	72.27	72.26
NSC+UPA	77.76	77.81	77.74	78.47	78.72	78.57
UPNN	81.02	80.67	80.21	79.29	80.28	79.77
ACNN	81.35	80.43	<b>80.85</b>	80.31	79.85	<b>80.02</b>
ASP-MTL	75.24	75.28	75.21	72.54	73.48	72.94
ACNN-AMT(full model)	81.90	80.73	<b>81.24</b>	82.43	81.74	<b>82.03</b>

Table 4: Comparison of microblog sentiment classification results.

language-independent sentiment features.

#### 5.4 Effect of Attention Mechanism and Multi-task Learning

Table 5 shows the effect of user attention in combination with pure CNN and multi-task CNN (CNN-MT) models. We can observe that: 1) The ACNN model and ACNN-MT model yield better sentiment classification results compared with CNN model and CNN-MT model without attention. It demonstrates that our user attention mechanism can capture crucial components for each user. 2) The CNN-MT model improves slightly compared with CNN model, and the reason is that there are substantial differences between Chinese and English languages, curbing from extracting global common sentiment features. The ACNN-MT model yields much better results than ACNN model, confirming that multi-task learning with user to bridge the gap between two languages can help generate more accurate user representation. We can also see that the ACNN-MT model improves more obviously on Weibo dataset. This is because larger proportion of Weibo posts than that of Twitter posts share users cross-lingually.

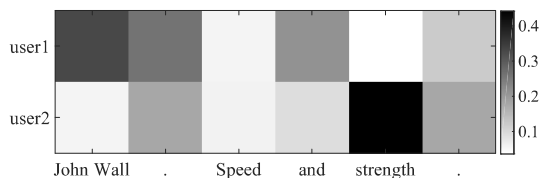


Figure 4: Visualization of attentions over unigram of Twitter. The darker color means higher attention weight.

In Figure 4 and Figure 5, we visualize the attention weights over  $n$ -grams to justify the effectiveness of our ACNN model. In Figure 4, we

choose a tweet posted by user1 with content “John Wall. Speed and strength”. The user1 is a fan of “John Wall”, and the user2 has never talked about him. For the word “John Wall”<sup>4</sup>, user1 has the largest attention weight, but user2 almost pays no attention, demonstrating that our ACNN model can capture crucial components for each user.

In Figure 5, we visualize the weights of a Chinese post posted by user1 with content “奇才 后卫 约翰沃尔 在美国 国家队 表演赛 一显身手” (“The Wizards’s guard John Wall displayed his skills in the exhibition match of US national team”). As shown in this post, user1’s pursuit of John Wall is unchanged in Chinese, indicating that analyzing texts jointly could help extract more precise individuality of a user. For unigram, user1 does not have the largest attention weight for word “约翰沃尔” (“John Wall”). Nonetheless, for bi-grams and tri-grams, user1 has the largest attention weights on “后卫 约翰沃尔” (“guard John Wall”) and “奇才 后卫 约翰沃尔” (“Wizards guard John Wall”), respectively, and the reason is that  $n$ -grams contain more abundant and accurate semantic information that is crucial for each user to compose appropriate post representation.

#### 5.5 Effect of Adversarial Mechanism

Table 6 shows the effect of adversarial mechanism. Chen et al. (2016b) has proved the important role of bilingual word embeddings in cross-lingual classification with adversarial learning. We train our bilingual word embeddings with MUSE (Conneau et al., 2017) so that semantically similar words in different languages are drawn closely in the embedding space. When the model is trying to classify Twitter posts, we transform Chi-

<sup>4</sup>The “John Wall” is the name of a basketball star, and we treat it as a single word.

Models	Twitter			Weibo		
	P	R	F	P	R	F
CNN	75.60	77.02	76.41	73.57	73.60	73.55
ACNN	81.35	80.43	80.85	80.31	79.85	80.02
CNN-MT	76.62	76.84	76.69	73.95	73.75	73.72
<b>ACNN-MT</b>	80.91	80.95	<b>80.89</b>	81.20	81.63	<b>81.53</b>

Table 5: Effect of attention mechanism and multi-task learning.

Models	Twitter			Weibo		
	P	R	F	P	R	F
CNN-MT	76.62	76.84	76.69	73.95	73.75	73.72
CNN-AMT	77.14	77.74	77.41	74.18	74.47	74.25
ACNN-MT	80.91	80.95	80.89	81.20	81.63	81.53
<b>ACNN-AMT</b>	81.90	80.73	<b>81.24</b>	82.43	81.74	<b>82.03</b>

Table 6: Effect of adversarial mechanism.

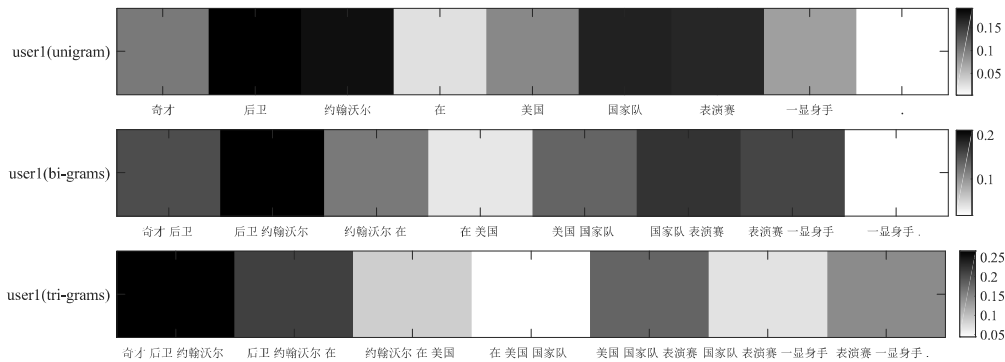


Figure 5: Visualization of attentions over  $n$ -grams of Weibo. (奇才 : Wizards; 后卫 : rearguard; 约翰沃尔: John Wall; 在: in; 美国: US; 国家队: national team; 表演赛: exhibition game; 一显身手: display his skills.)

nese word embeddings into English word embeddings for alignment, and vice versa. By comparing models with adversarial mechanism (i.e., CNN-AMTL, ACNN-AMTL) and models without adversarial mechanism (i.e., CNN-MTL, ACNN-MTL), the former yield better results, suggesting that the adversarial learning could help generate language-independent sentiment features, and improve the effect of sentiment classification.

## 6 Conclusion

In this paper, we proposed an ACNN model for capturing user’s individuality on microblogs, and an ACNN-AMT model for alleviating the data insufficiency problem in personalized sentiment classification. By exploiting the user information, our ACNN model could put emphasis on the key portions of posts that reflect users’ individu-

ality. By utilizing the consistent individuality and opinion bias of microblog users across different languages, our ACNN-AMT model could extract the language-specific features and language-independent features to enrich the user-specific post representation. Experimental results on English and Chinese microblog datasets confirm the clear advantage of our method over state-of-the-art baseline algorithms.

## Acknowledgments

The work was supported by the National Key R&D Program of China under grant 2018YF-B1004700, and National Natural Science Foundation of China (61772122, 61872074).

## References

- Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.
- Chih-Chung Chang and Chih-Jen Lin. 2011. Libsvm: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3):27.
- Huimin Chen, Maosong Sun, Cunchao Tu, Yankai Lin, and Zhiyuan Liu. 2016a. Neural sentiment classification with user and product attention. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1650–1659.
- Xilun Chen, Yu Sun, Ben Athiwaratkun, Claire Cardie, and Kilian Weinberger. 2016b. Adversarial deep averaging networks for cross-lingual sentiment classification. *arXiv preprint arXiv:1606.01614*.
- Alexis Conneau, Guillaume Lample, Marc’Aurelio Ranzato, Ludovic Denoyer, and Hervé Jégou. 2017. Word translation without parallel data. *arXiv preprint arXiv:1710.04087*.
- Theodoros Evgeniou and Massimiliano Pontil. 2004. Regularized multi-task learning. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 109–117. ACM.
- Yaroslav Ganin and Victor Lempitsky. 2015. Unsupervised domain adaptation by backpropagation. In *International Conference on Machine Learning*, pages 1180–1189.
- Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680.
- Vibhu Jawa and Varun Hasija. 2015. A sentiment and interest based approach for product recommendation. In *Modelling and Simulation (UKSim), 2015 17th UKSim-AMSS International Conference on*, pages 75–80. IEEE.
- Nal Kalchbrenner, Edward Grefenstette, and Phil Blunsom. 2014. A convolutional neural network for modelling sentences. *arXiv preprint arXiv:1404.2188*.
- Yoon Kim. 2014. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*.
- Bing Liu. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167.
- Pengfei Liu, Xipeng Qiu, and Xuanjing Huang. 2017. Adversarial multi-task learning for text classification. *arXiv preprint arXiv:1704.05742*.
- Christopher Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard, and David McClosky. 2014. The stanford corenlp natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*, pages 55–60.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543.
- Kaisong Song, Ling Chen, Wei Gao, Shi Feng, Daling Wang, and Chengqi Zhang. 2016. Persentiment: A personalized sentiment classification system for microblog users. In *Proceedings of the 25th International Conference Companion on World Wide Web*, pages 255–258. ACM.
- Kaisong Song, Shi Feng, Wei Gao, Daling Wang, Ling Chen, and Chengqi Zhang. 2015a. Build emotion lexicon from microblogs by combining effects of seed words and emoticons in a heterogeneous graph. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, pages 283–292. ACM.
- Kaisong Song, Shi Feng, Wei Gao, Daling Wang, Ge Yu, and Kam-Fai Wong. 2015b. Personalized sentiment classification based on latent individuality of microblog users. In *IJCAI*, pages 2277–2283.
- Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958.
- Xiao Sun, Fei Gao, Chengcheng Li, and Fuji Ren. 2015. Chinese microblog sentiment classification based on convolution neural network with content extension method. In *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*, pages 408–414. IEEE.
- Duyu Tang, Bing Qin, and Ting Liu. 2015. Learning semantic representations of users and products for document level sentiment classification. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1, pages 1014–1023.
- Fangzhao Wu and Yongfeng Huang. 2016. Personalized microblog sentiment classification via multi-task learning. In *AAAI*, pages 3059–3065.
- Fangzhao Wu, Yangqiu Song, and Yongfeng Huang. 2015. Microblog sentiment classification with contextual knowledge regularization. In *AAAI*, pages 2332–2338.

Zhiwei Yu, Raymond K Wong, Chi-Hung Chi, and Fang Chen. 2015. A semi-supervised learning approach for microblog sentiment classification. In *Smart City/SocialCom/SustainCom (SmartCity), 2015 IEEE International Conference on*, pages 339–344. IEEE.

Matthew D Zeiler. 2012. Adadelta: an adaptive learning rate method. *arXiv preprint arXiv:1212.5701*.

Zhou Zhao, Hanqing Lu, Deng Cai, Xiaofei He, and Yueting Zhuang. 2017. Microblog sentiment classification via recurrent random walk network learning. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pages 3532–3538. AAAI Press.