Emoji-Based Sentiment Analysis Using Attention Networks

YINXIA LOU, Wuhan University
YUE ZHANG, Westlake University
FEI LI, University of Massachusetts Lowell
TAO QIAN, Hubei University of Science and Technology
DONGHONG JI, Wuhan University

Emojis are frequently used to express moods, emotions, and feelings in social media. There has been much research on emojis and sentiments. However, existing methods mainly face two limitations. First, they treat emojis as binary indicator features and rely on handcrafted features for emoji-based sentiment analysis. Second, they consider the sentiment of emojis and texts separately, not fully exploring the impact of emojis on the sentiment polarity of texts. In this article, we investigate a sentiment analysis model based on bidirectional long short-term memory, and the model has two advantages compared with the existing work. First, it does not need feature engineering. Second, it utilizes the attention approach to model the impact of emojis on text. An evaluation on 10,042 manually labeled Sina Weibo showed that our model achieves much better performance compared with several strong baselines. To facilitate the related research, our corpus will be publicly available at https://github.com/yx100/emoji.

CCS Concepts: • **Information systems** → **Sentiment analysis**;

Additional Key Words and Phrases: Sentiment analysis, social media, emoji, deep learning, attention

ACM Reference format:
Yinxia Lou, Yue Zhang, Fei Li, Tao Qian, and Donghong Ji. 2020. Emoji-Based Sentiment Analysis Using Attention Networks. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* 19, 5, Article 64 (May 2020), 13 pages.
https://doi.org/10.1145/3389035

1 INTRODUCTION

Microblogging allows millions to express their feelings, emotions, and attitudes. Rich information is contained in microblog posts, such as emojis, hashtags, and videos, which makes them a hot

This work was supported by the National Natural Science Foundation of China (no. 61772378), the Major Projects of the National Social Science Foundation of China (no. 11&ZD189), the Natural Science Foundation of Hubei Province, China (no. 2018CFB690), and the Science and Technology Project of Guangzhou (no. 201704030002).
Authors' addresses: Y. Lou, Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, School of Cyber Science and Engineering, Wuhan University, 299 Bayi Road, Wuhan City, 430072, China; email: yinxia@whu.edu.cn; Y. Zhang, Westlake University, 18 Shilongshan Road, Hangzhou City, 310024, China; email: yue.zhang@wias.org.cn; F. Li, University of Massachusetts Lowell, 1 University Ave, Lowell, MA 01854, USA; email: foxlf823@gmail.com; T. Qian, Hubei University of Science and Technology, 30 Shuangqing Road, Xianning City, 437100, China; email: taoqian@whu.edu.cn; D. Ji (corresponding author), Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, School of Cyber Science and Engineering, Wuhan University, 299 Bayi Road, Wuhan City, 430072, China; email: dhji@whu.edu.cn.

2375-4699/2020/05-ART64 $15.00
https://doi.org/10.1145/3389035

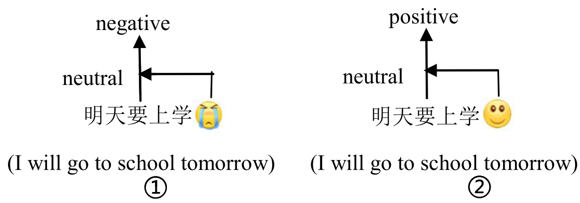

Fig. 1. The impact of emojis on the sentiment polarity of text.

research target. In particular, emojis are becoming increasing popular [32, 43] and have been investigated in sentiment analysis, one of the most basic tasks and key topics in microblog research [9, 20, 23, 34, 37, 49]. The purpose is to automatically analyze the polarity of a microblog post, which can be positive, negative, or neutral [2, 4, 5, 8, 15, 44].

Emojis and sentiments have attracted attention in both sociology and computer science. Sociology research uses statistical methods to analyze the intentions between emoji usage and sentiment effects of emojis in a microblog post [14, 40, 45]. Computer science research investigates models to predict the sentiment polarity of a microblog post with emojis. Previous work mainly used emojis as features, among other designed features, to improve the performance of sentiment analysis. For example, Mohammad et al. [29] used rich linguistically motivated features from tweets for sentiment analysis. They used not only simple features such as emojis, lexical features (word n-grams, character n-grams, and elongated words), lexicons, and punctuation features but also sophisticated features such as part-of-speech (POS) tags and Brown clusters.

Current work on emojis mainly faces two limitations. First, they rely on manual indicator features, which can be sparse and weak for semantic representation. Second, they consider the sentiments of emojis and plain texts separately, not fully exploring the impact of emojis on the sentiment polarity of texts. Emojis play an important role in the sentiment polarity of plain texts. As an example, Figure 1 shows the impact of emojis on the sentiment polarity of texts, where the sentiment of the plain text is originally neutral. If the text is augmented with 😭 or 🙂 in the end, the posts convey totally different sentiment polarities, namely negative and positive. In this work, we aim at investigating the impact of emojis on the sentiment polarity of texts to predict the sentiment polarity of the microblog post as a whole.

We propose a deep learning architecture to model the impact of emojis on the sentiment polarity of text for sentiment analysis. As illustrated in Figure 2, our model mainly consists of three parts. First, we build bidirectional long short-term memory (Bi-LSTM) to capture the representation of a microblog post. Second, to obtain the impact of emojis on the sentiment polarity of text, we use attention [41] to weigh each word based on the emoji. Finally, we concatenate the text representation, emoji representation, and emoji-weighted text representation as the input of the sentiment analysis model for predicting the sentiment polarity of a post.

Although there have been some annotated corpora on Chinese and English for sentiment analysis, such as SemEval2015 [36], SemEval2016 [30], and MVSC [33], they do not explicitly model the interaction between emojis and text. To fill this gap, we manually annotate a Chinese microblog corpus, which contains the polarities of microblog posts with and without emojis. Experimental results show the effctiveness of our model compared with several strong baselines, including traditional shallow learning and neural network models.

The main contributions of our work can be summarized as follows:

- We build and release a Chinese microblog corpus with emojis, which contains 10,042 microblog posts. This corpus considers the impacts of both text and emojis on the sentiment polarity.

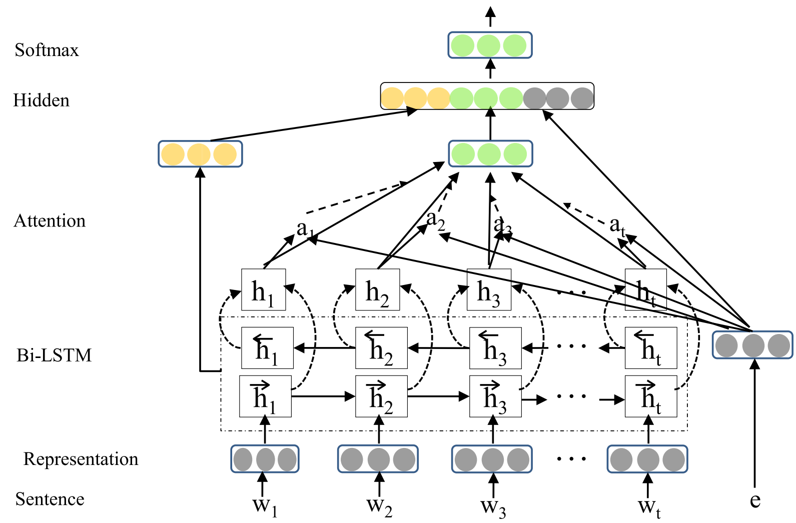

Fig. 2. The architecture of an emoji attention-based neural sentiment analysis model.

- We jointly train emojis and words in microblog posts and obtain the emoji representations containing their contextual information.
- To the best of our knowledge, we are the first to propose an attention model to capture the impact of emojis on the sentiment polarity of text.

2 RELATED WORK

Sentiment analysis [23, 25, 44] has attracted much attention in the domain of natural language processing (NLP). Emojis are "picture characters" or pictographs that began to appear on mobile phones in the late 1990s. Recently, emojis have replaced emoticons and have been widely adopted for simplifying the expression of emotions and enriching the communications on social media, such as Sina Weibo, Twitter, and Facebook [1, 17, 21, 32, 40].

Seminal work used emojis as noisy sentiment labels to train classifiers [11, 31]. Go et al. [11] identified the tweet polarity using emojis as noisy labels and collected a training dataset of 1.6 million tweets. However, the performance of such models can be limited due to noise in the labels.

With the development of NLP, most prior work mainly focused on designing effective features to improve the sentiment classification performance [10, 29]. For example, Mohammad et al. [29] constructed SVM classifiers with sparse indicator features including n-grams, POS tags, punctuations, emojis, and clusters. In contrast to linear models such as SVM, neural network models automatically extract features and have achieved promising results for sentiment classification [1, 22, 25, 35]. Tang et al. [39] introduced a neural network model to learn vector-based document representation for document-level sentiment classification. Kim [19] used convolutional neural network (CNN) models for sentence-level classification tasks. Most similar to our motivation, Le et al. [22] proposed LSTMs to analyze sentiment on Indonesian tweets and obtained promising results. They first translated emojis into their equivalent words and then obtained the embeddings of these words. Their method outperforms traditional shallow learning algorithms. Although they used real-valued word embeddings to solve the feature sparsity problem of discrete models, their model treats emojis and text in a microblog post as two separate parts, without explicitly considering the impact of emojis on the sentiment polarity of text. In contrast, we capture long distance sentiment dependency in microblog posts using Bi-LSTM models and consider the impact of emojis on the sentiment polarity of text.

In fact, the emojis in microblogs have effects on sentiment polarity. Sociology research has found evidence of this phenomenon [3, 14, 40]. However, the research mainly analyzed typical intentions of emojis in communication and the sentiment effects of emojis from a sociological perspective and did not study this from the point of computational linguistics. In contrast, we design an emoji-based attention mechanism to capture the effects. The attention is to select crucial words from the whole word sequence in a microblog post.

Previous studies have shown that the attention mechanism can be effectively used in many tasks of NLP, such as machine translation [27], parsing [24, 42], document classification [46], text understanding [18], and question answering systems [38]. Attention has been applied for sentiment analysis [26], such as the aspect sentiment [7], user-oriented sentiment [6], and cross-lingual sentiment [50]. To the best of our knowledge, we are the first to use attention to model the impact of emojis on the sentiment polarity of text for sentiment analysis.

3 DATASET CREATION

Existing corpora of sentiment analysis contain only a small fraction with emojis. These corpora are not particularly suitable for emoji-based sentiment analysis. We describe the process of collecting and annotating microblog posts with emojis, including the text polarity and the overall polarity of microblog posts with emojis.

3.1 Data Collection

We collected 300,000 microblog posts from the Sina Weibo website,[1] which is one of the most popular microblog sites in China. Then, we extracted 110,000 microblog posts that contained emojis. We ranked microblog posts according to the occurrence of each emoji and selected the set of emojis that occurred at least 10 times. Finally, we split each microblog post by emojis and selected microblog posts with only one emoji. We filtered out URLs, user names, and hashtags to clean the data. Microblog posts with lengths greater than 5 were retained. Then, 80,000 microblog posts were left. Finally, we randomly took 15,000 microblog posts for labeling in the next step. The Jieba Chinese text segmentation tool[2] was used for segmentation.

3.2 Annotation

We hired three annotators to construct this corpus: one senior linguistics student and two students majoring in computer science. Sentiment polarities were classified into positive, neutral, and negative, denoted by 0, 1, and 2, respectively. A marked label appearing at least twice would be accepted.

The annotation work was mainly divided into two parts. First, annotators were asked to label the polarity of each post based only on text. In other words, emojis were removed from the text and only the plain text of each microblog post was used as the evidence of the polarity. Second, annotators were asked to label each post by considering both text and emojis. We finally labeled the polarities of 10,042 microblog posts with emojis. Table 1 shows the corpus statistics, where column 5 is interannotator consistency of three labels.

[1] https://weibo.com.
[2] https://github.com/fxsjy/jieba.

Table 1. Corpus Statistic with Row 1 and Row 2 Denoting Polarity of Plain Texts and Microblogs with Emojis, Respectively

Corpus	Positive	Neutral	Negative	Consistency
Text polarity	3,827 (38%)	3,618 (36%)	2,597 (26%)	85%
Overall polarity	5,819 (58%)	902 (9%)	3,321 (33%)	72%

Table 2. Statistics of Nonchanges and Changes in Polarities of Microblog Posts

Sentiment	Polarity		Microblogs
	Text	Overall	
Nonchanges	Positive	Positive	3,556
	Neutral	Neutral	334
	Negative	Negative	2111
	Total		6,001
Changes	Positive	Neutral	180
	Positive	Negative	91
	Neutral	Positive	2,162
	Neutral	Negative	1,119
	Negative	Positive	101
	Negative	Neutral	388
	Total		4,041

Text denotes microblog polarities without emojis. *Overall* denotes microblog polarities with emojis.

3.3 Corpus Analysis

Emojis may change the sentiment polarities of microblog posts by subtle interaction with text. We investigated microblog posts whose sentiment polarities were changed and unchanged, as shown in Table 2. There were 4,044 microblog posts whose polarities changed under the effects of emojis, accounting for 40.27% of all microblog posts.

4 MODEL

An overview of our model is shown in Figure 2. In this section, we introduce our neural sentiment analysis (NSA) model based on emoji attention (EA). First, we explain how to obtain the text semantic representation via the Bi-LSTM network. Then, our EA approach is introduced. Last, we describe the training process of our EA-Bi-LSTM model.

4.1 Bi-LSTM-Based Sentiment Analysis Model

Bi-LSTM is a variation of the recurrent neural network (RNN) [12], which has been widely used in NLP. In sentiment analysis, the Bi-LSTM model is applied to learn the representation of a sentence, then the representation is used as features to classify the sentiment. Yang et al. [47] applied a Bi-LSTM model to text classification and achieved excellent performance.

LSTM is used to capture long range dependencies in sequences [13]. An LSTM model has multiple LSTM cells, where each LSTM cell models the memory in a neural network. It has several gates that allow the LSTM to store and access information over time. Given a short text with words w_t, $t \in [1, T]$, the words are embedded to their vectors through an embedding matrix W_e, $x_t = W_e w_t$,

$x_t \in \mathbb{R}^d$, where d is the dimension of word embeddings. Our model adopts Bi-LSTM for reading text bidirectionally. Bi-LSTM contains a forward $\overrightarrow{LSTM}$ that reads the text from x_1 to x_T and a backward $\overleftarrow{LSTM}$ that reads the text from x_T to x_1, formalized by

$$\begin{aligned} \overrightarrow{h_t} &= \overrightarrow{LSTM}(x_t), t \in [1, T], \\ \overleftarrow{h_t} &= \overleftarrow{LSTM}(x_t), t \in [1, T]. \end{aligned} \tag{1}$$

Bi-LSTM maps each word w_t to a pair of hidden vectors $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$, so a word can be represented as the concatenation $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$, formalized by $h_t = [\overrightarrow{h_t}, \overleftarrow{h_t}]$. Therefore, we get $[h_0, h_1, h_2, \ldots, h_T]$ and then feed them to an average pooling layer to obtain a sentence representation s.

4.2 Attention for Emoji-Based Sentiment Analysis Model

The process of sentiment change is similar to the attention mechanism in that useful information is selected in text [16]. To indicate the impact of emojis on the sentiment polarity of text, we propose an emoji-based attention mechanism. Given a microblog post, each word contributes unequally to the sentiment polarity, and the interaction weights of emojis are also unequal. The EA mechanism measures the weights of words in a microblog post after incorporating words and emojis.

In a microblog post $\{w_1, w_2, \ldots, w_T; E\}$, w_i denotes the token and E denotes the emoji. First, both w_i and E are converted to vector representations, namely $x_i \in \mathbb{R}^d$ and $e \in \mathbb{R}^d$, where d is the dimension of the vector.

Different from the preceding section, $[h_1, h_2, \ldots, h_T]$ are denoted as the representations of the text $\{w_1, w_2, \ldots, w_T\}$ by the Bi-LSTM layer. We aggregate the representations of those informative words to form the sentence representation. A sentence representation s is computed as a weighted sum of the hidden state h_i of its word as

$$s = \sum_{i=1}^{T} a_i h_i, \tag{2}$$

where a_i measures the importance of the i-th word. The attention weight a_i for each hidden state can be defined as

$$a_i = \frac{\exp(score(h_i, e))}{\sum_{j=1}^{T} \exp(score(h_j, e))}, \tag{3}$$

where $score$ indicates the importance of words. The $score$ is defined as

$$score(h_i, e) = v^T \tanh(W_h h_i + W_E e + b), \tag{4}$$

where $W_h, W_E \in \mathbb{R}^{a \times d}$, and $v \in \mathbb{R}^a$ are learnable parameters; v^T denotes the transpose of v; and b is the bias. Finally, we concatenated three types of features:

$$l_c = [\overrightarrow{h_0}, \overleftarrow{h_T}] \oplus s \oplus e, \tag{5}$$

where $\overrightarrow{h_0}$ and $\overleftarrow{h_T}$ represent the hidden states of the forward and backward LSTMs in the last step.

4.3 Training

Our training objective is to minimize the cross-entropy loss. After introducing the emoji-based attention mechanism, we obtained final features l_c for sentiment analysis of the text. Our model uses a linear transformation to project l_c into the target space of C classes:

$$d_c = W_c l_c + b_c. \tag{6}$$

Afterward, we used a softmax layer to obtain the probability distribution of the microblog post sentiment:

$$p_c = \frac{\exp(d_c)}{\sum_{k=1}^{C} \exp(d_k)}, \tag{7}$$

where C is the number of sentiment labels and p_c is the predicted probability for the sentiment label c.

Let $p_c^g(d)$ be the target distribution for a post, $p_c(d)$ be the predicted sentiment distribution, and D be the set of microblog posts. The training objective is to minimize the cross-entropy loss between $p_c^g(d)$ and $p_c(d)$ for D. The loss function is defined as

$$L = -\sum_{d \in D} \sum_{c=1}^{C} p_c^g(d) \log(p_c(d)). \tag{8}$$

5 EXPERIMENTS

In this section, we first describe our experimental settings. Then, we introduce several baseline models including the state-of-the-art method for comparisons. Finally, we introduce the empirical results with corresponding discussions.

5.1 Experimental Settings

5.1.1 Embeddings. To obtain the embedding representations of words and emojis in microblogs, a word or an emoji embedding was trained on a large-scale corpus consisting of 3.5 million Chinese microblogs. Words and emojis were trained simultaneously using the SkipGram mode [28] of word2vec.[3] The vocabulary size was 252,267. We randomly initialized word or emoji embeddings that were out of vocabulary and performed supervised fine tuning over the training corpus.

5.1.2 Evaluation Method. We used fivefold cross validation in our experiments. Typically, original data were randomly split into five equal sections, where four sections were selected for training and the fifth section was used for testing. We randomly chose one section from the four training sections as the development set to tune hyperparameters. The classification results were measured by accuracy, defined as

$$Accuracy = \frac{T}{N}, \tag{9}$$

where T indicates the number of predicted sentiment ratings that are identical with gold sentiment ratings and N indicates the number of microblogs. Due to the class imbalance problem in multiclassification, we also used macroaccuracy for a fairer comparison.

5.1.3 Hyperparameters. We set the dimensions of word embeddings and emoji embeddings as 200. The dimensions of hidden states and cell states in our LSTM cells were set to 100. We used Adadelta [48] as our optimization method during training. We trained all models with the batch size of 16, the momentum as 0.9, and the initial learning rate α as 0.01.

5.2 Baselines

To evaluate the performance of our EA-Bi-LSTM model, we compared it with several baselines, including EMOJI-Noisy labels, EMOJI-EMB, SVM, LSTM (text+emoji), Bi-LSTM (text), Bi-LSTM

[3] https://code.google.com/p/word2vec.

(text+emoji), and EA-Bi-GRU (text+emoji). SVM and LSTM (text+emoji) sentiment analysis models were reimplemented on our dataset. Further details of the datasets include the following:

- *EMOJI-Noisy labels*: We used emojis as noisy labels and directly computed the accuracy of labels using the following formula: the correct number of microblog posts labeled by emojis/the total number of microblog posts.
- *EMOJI-EMB* [22]: We used only emoji embedding to predict the sentiment polarity of a microblog post.
- *SVM* [29]: A statistical method for binary classification, which does not take the impact of emojis on the sentiment polarity of text into account. To train the classifier, we used features such as emojis, bag-of-words, and punctuation.
- *LSTM (text+emoji)* [22]: LSTM was used for sentiment analysis. This model learns the vector representations of words and emojis from microblog posts.
- *Bi-LSTM (text)*: We used only plain text of microblogs as the inputs to the Bi-LSTM model for sentiment analysis.
- *Bi-LSTM (text+emoji)*: We took both the text and emojis of microblog posts as input to the Bi-LSTM model for sentiment analysis.
- *EA-Bi-GRU (text+emoji)*: We used GRU instead of LSTM cells in the EA-Bi-LSTM model to verify the effectiveness of LSTM for short texts.

5.3 Results

Table 3 shows the experiment results of all models for sentiment analysis on the Chinese Sina microblog corpus. Because of the class imbalance problem, the performance of neutral microblogs is much lower than those of positive and negative microblogs. To evaluate our model fairly, we used two types of measures—accuracy and macroaccuracy, which achieved consistent performance on our corpus.

In Table 3, we see that the EMOJI-Noise labels and EMOJI-EMB models improve the accuracy by 15.71% and 15.98%, respectively, compared with the Bi-LSTM (text) model. It demonstrates that the impact of emojis on sentiment polarity of a microblog post is stronger than that of text, which can also be confirmed by the results of the Bi-LSTM (text+emoji), being higher than the Bi-LSTM (text) model. Furthermore, we found that the models only using an emoji feature were not better than those neural network models using both emoji and text features. The best model, Bi-LSTM (text+emoji) using two features, outperformed the EMOJI-EMB model by 1.10% and 3.47% in accuracy and macroaccuracy, respectively. This shows that both text and emojis play important roles in sentiment prediction of microblog posts.

Comparing the LSTM (text+emoji) neural network with the discrete model SVM, experimental results show that LSTM (text+emoji) outperforms the SVM. This demonstrates that neural network models are a strong choice for extracting text and emoji features compared to the discrete models with sparse indicator features.

The results in Table 3 show that our EA-Bi-LSTM model performs the best and significantly outperforms all baselines. The performance of the EA-Bi-GRU model was slightly worse than that of the EA-Bi-LSTM model, which shows that LSTM is a reasonable choice for the short text setting. The EA-Bi-LSTM model achieved 1.10% accuracy improvement and 3.47% macroaccuracy improvement over Bi-LSTM (text+emoji), respectively. Compared with the Bi-LSTM (text+emoji) that uses two features, our EA-Bi-LSTM model utilizes features including text, emojis, and the impact of emojis on text. This demonstrates that emoji-based attention can effectively capture the impact of emojis on the sentiment polarity of text. We also used precision (P), recall (R), and F-score (F) as our assessing criteria to evaluate our model in Table 3. As we can see, the EA-Bi-LSTM model also

Table 3. Results of Different Models

Models	Polarity	P (%)	R (%)	F (%)	Acc (%)	Macro-Acc (%)
EMOJI-Noise labels	—	—	—	—	85.49	—
EMOJI-EMB	Positive	88.42	92.90	90.64	85.76	65.04
	Neutral	37.80	18.73	20.19		
	Negative	88.25	90.15	89.32		
SVM	Positive	81.82	83.76	82.78	78.54	61.76
	Neutral	36.00	22.91	28.00		
	Negative	79.00	82.93	80.92		
LSTM (text+emoji)	Positive	88.51	94.08	89.79	86.16	65.70
	Neutral	38.60	18.29	21.86		
	Negative	88.34	90.87	89.55		
Bi-LSTM (text)	Positive	74.21	83.48	78.52	69.78	50.71
	Neutral	18.95	4.34	6.64		
	Negative	64.13	63.40	63.54		
Bi-LSTM (text+emoji)	Positive	87.22	95.34	90.69	86.66	68.47
	Neutral	43.14	15.83	22.33		
	Negative	90.53	90.85	90.49		
EA-Bi-GRU	Positive	89.71	92.05	90.86	87.01	69.24
	Neutral	39.73	28.86	32.51		
	Negative	89.88	90.86	90.34		
EA-Bi-LSTM	Positive	89.23	94.43	**91.71**	**87.85**	**69.80**
	Neutral	46.89	26.68	**33.37**		
	Negative	91.26	91.69	**91.45**		

P, *R*, and *F* denote precision, recall, and F-score, respectively. *Acc* denotes accuracy, and *Macro-Acc* denotes macroaccuracy.

achieves the best performance in terms of F-scores, which are 91.72%, 33.37%, and 91.45% on three sentiment polarities, respectively.

5.4 Analysis

The impact of emojis on the sentiment polarity of text. We selected the experimental results of the Bi-LSTM (text+emoji) model and the EA-Bi-LSTM model, respectively, analyzing the accuracies of two models for different sentiment polarities. Moreover, we analyzed the performance of two models in Table 2. Table 4 shows the accuracies of the two models for different sentiment polarities.

From Table 4, we can see that our EA-Bi-LSTM model improved the accuracies of neutral and negative sentiment by 7.79% and 0.67% compared with the Bi-LSTM (text+emoji) model in the aspect of nonchange sentiment polarities. The mean of the EA-Bi-LSTM model was also 3.01% higher than that of the Bi-LSTM (text+emoji) model without changing the sentiment.

In terms of sentiment change, our EA-Bi-LSTM model outperformed the Bi-LSTM (text) model in most cases. Especially, our EA-Bi-LSTM model significantly improved the accuracies by 5.56%, 3.29%, and 7.73% in the cases where polarities change from positive to neutral, from positive to negative, and from negative to neutral, respectively. This demonstrates that our model can make better use of the effects of emojis on text for sentiment analysis.

5.5 Case Study

To show the difference between our EA-Bi-LSTM model and the Bi-LSTM (text+emoji) model, we randomly sampled some examples as shown in Figure 3. Columns 2 through 4 represent the

Table 4. Results of Different Sentiment Polarities

Sentiment	Polarity		Bi-LSTM	EA-Bi-LSTM
	Text	Overall	(text+emoji) Acc (%)	Acc (%)
Nonchanges	Positive	Positive	95.24	95.84
	Neutral	Neutral	27.54	35.33
	Negative	Negative	88.58	89.25
	Average		70.46	**73.47**
Changes	Positive	Neutral	11.11	16.67
	Positive	Negative	84.62	87.91
	Neutral	Positive	95.56	97.80
	Neutral	Negative	95.00	97.05
	Negative	Positive	86.14	89.11
	Negative	Neutral	10.31	18.04
	average		63.79	**67.76**

Columns 4 and 5 represent the accuracies (Acc) of Bi-LSTM (text+emoji) and EA-Bi-LSTM, respectively. *Text* denotes microblog polarities without emojis. *Overall* denotes microblog polarities with emojis.

Microblogs	gold polarity	Bi-LSTM (text +emoji)	EA-Bi-LSTM
1) 快 抓 快 抓！ 别 让 它们 飞 了！ (Grab fast! Don't let them fly!)	positive	negative	positive
2) 又转 桌面 了 (Turned to the desktop again)	negative	positive	negative
3) ？ 钱 真是 万能 的！ (? Money is really omnipotent!)	negative	neutral	negative
4) 明天 早上 六点 街道口 地下通道 开始 通车， 街道口 不会 在 堵 了！? 大家 终于 等到 这 一天 啦 (At six tomorrow morning, the underground passage of the street entrance will start to open, and the street entrance will not be blocked !? Everyone finally waited until this day.)	positive	neutral	positive

Fig. 3. Microblog samples of EA-Bi-LSTM predict right, but Bi-LSTM (text+emoji) predicts wrong.

gold polarity, predicted polarity by Bi-LSTM (text+emoji), and predicted polarity by EA-Bi-LSTM, respectively. We can see that Bi-LSTM (text+emoji) gives incorrect predictions in all of these examples, whereas our model performs well. One likely reason is that Bi-LSTM (text+emoji) equally treats emojis and text, but our model pays attention to only important words and emojis.

It can be enlightening to analyze which word decides the sentiment polarity of the microblog considering the emoji. We can obtain the attention weight a in Equation (5) and visualize the attention weights accordingly. Figure 4 shows how attention helps modeling the importance of a word with respect to the emoji in a microblog. We use a histogram to represent the weight of each word. The vertical axis indicates the weight of each word, and the horizontal axis represents words in a microblog text. The column height indicates the importance of the word. As shown in Figure 4, the word "可惜 (pity)" has the highest score, indicating that it can play an important role in analyzing sentiment of the whole sentence.

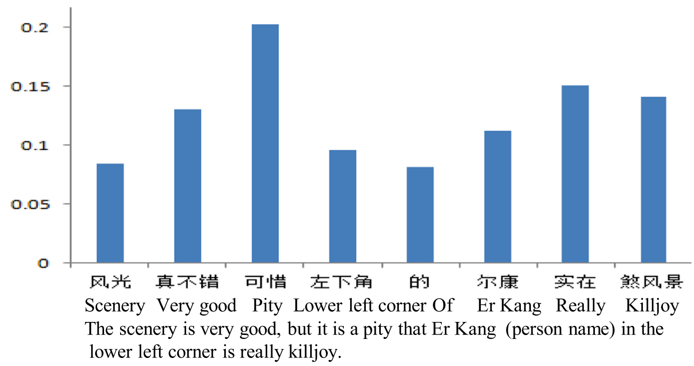

Fig. 4. Attention visualizations.

6 CONCLUSION AND FUTURE WORK

We have proposed an attention model to improve emoji-based sentiment analysis on microblog posts. Our model takes full advantage of the impact of emojis on the sentiment polarity of texts. We simultaneously trained emoji and text embeddings. Compared with several strong baseline models, our model achieves the highest performance. Moreover, we constructed a large-scale annotated corpus of a Chinese microblog that contains both plain text polarities and text-emoji polarities. To the best of our knowledge, we are the first to use an attention mechanism to model the impact of emojis on the sentiment polarity of texts. In the future, we will further study the effect of emojis on the sentiment polarity of short texts in two directions. First, we will extend the research to other types of short texts, such as tweets and WeChat. Second, we will investigate more neural network models, such as joint models or multitask learning models, to explore the impact of emojis on texts.

REFERENCES

[1] Sadam Al-Azani and El-Sayed El-Alfy. 2018. Emojis-based sentiment classification of Arabic microblogs using deep recurrent neural networks. In *Proceedings of the International Conference on Computing Sciences and Engineering (ICCSE'18)*. IEEE, Los Alamitos, CA, 1–6.

[2] Francesco Barbieri, Luis Espinosa Anke, Jose Camacho-Collados, Steven Schockaert, and Horacio Saggion. 2018. Interpretable emoji prediction via label-wise attention LSTMs. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 4766–4771.

[3] Naomi S. Baron. 2009. The myth of impoverished signal: Dispelling the spoken language fallacy for emoticons in online communication. In *Electronic Emotion: The Mediation of Emotion via Information and Communication Technologies*. Peter Lang, 107–135.

[4] Ghazaleh Beigi, Xia Hu, Ross Maciejewski, and Huan Liu. 2016. An overview of sentiment analysis in social media and its applications in disaster relief. In *Sentiment Analysis and Ontology Engineering*. Springer, 313–340.

[5] Johan Bollen, Huina Mao, and Alberto Pepe. 2011. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In*Proceedings of the 5th International AAAI Conference on Weblogs and Social Media (ICWSM'11)*. 450–453.

[6] Huimin Chen, Maosong Sun, Cunchao Tu, Yankai Lin, and Zhiyuan Liu. 2016. Neural sentiment classification with user and product attention. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. 1650–1659.

[7] Peng Chen, Zhongqian Sun, Lidong Bing, and Wei Yang. 2017. Recurrent attention network on memory for aspect sentiment analysis. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 452–461.

[8] Yuxiao Chen, Jianbo Yuan, Quanzeng You, and Jiebo Luo. 2018. Twitter sentiment analysis via bi-sense emoji embedding and attention-based LSTM. In *Proceedings of the 2018 ACM Conference on Multimedia (CM'18)*. 117–125.

[9] Zhenpeng Chen, Sheng Shen, Ziniu Hu, Xuan Lu, Qiaozhu Mei, and Xuanzhe Liu. 2019. Emoji-powered representation learning for cross-lingual sentiment classification. In *Proceedings of the World Wide Web Conference*. ACM, New York, NY, 251–262.

[10] Xing Fang and Justin Zhan. 2015. Sentiment analysis using product review data. *Journal of Big Data* 2, 1 (2015), 5.

[11] Alec Go, Richa Bhayani, and Lei Huang. 2009. *Twitter Sentiment Classification Using Distant Supervision*. CS224N Project Report. Stanford University.

[12] Alex Graves and Jürgen Schmidhuber. 2005. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks* 18, 5–6 (2005), 602–610.

[13] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation* 9, 8 (1997), 1735–1780.

[14] Tianran Hu, Han Guo, Hao Sun, Thuy-Vy Thi Nguyen, and Jiebo Luo. 2017. Spice up your chat: The intentions and sentiment effects of using emoji. arXiv:1703.02860.

[15] Xia Hu, Lei Tang, Jiliang Tang, and Huan Liu. 2013. Exploiting social relations for sentiment analysis in microblogging. In *Proceedings of the 6th ACM International Conference on Web Search and Data Mining*. ACM, New York, NY, 537–546.

[16] Fei Jiang, Yiqun Liu, Huanbo Luan, Min Zhang, and Shaoping Ma. 2014. Microblog sentiment analysis with emoticon space model. In *Proceedings of the Chinese National Conference on Social Media Processing*. 76–87.

[17] Fei Jiang, Yi-Qun Liu, Huan-Bo Luan, Jia-Shen Sun, Xuan Zhu, Min Zhang, and Shao-Ping Ma. 2015. Microblog sentiment analysis with emoticon space model. *Journal of Computer Science and Technology* 30, 5 (2015), 1120–1129.

[18] Rudolf Kadlec, Martin Schmid, Ondrej Bajgar, and Jan Kleindienst. 2016. Text understanding with the attention sum reader network. arXiv:1603.01547.

[19] Yoon Kim. 2014. Convolutional neural networks for sentence classification. arXiv:1408.5882.

[20] Mayu Kimura and Marie Katsurai. 2017. Automatic construction of an emoji sentiment lexicon. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. ACM, New York, NY, 1033–1036.

[21] Svetlana Kiritchenko, Xiaodan Zhu, and Saif M. Mohammad. 2014. Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research* 50 (2014), 723–762.

[22] Tuan Anh Le, David Moeljadi, Yasuhide Miura, and Tomoko Ohkuma. 2016. Sentiment analysis for low resource languages: A study on informal Indonesian tweets. In *Proceedings of the 12th Workshop on Asian Language Resources (ALR-12)*. 123–131.

[23] Bing Liu. 2012. *Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies.* Claypool.

[24] Jiangming Liu and Yue Zhang. 2016. Shift-reduce constituent parsing with neural lookahead features. arXiv:1612.00567.

[25] Jiangming Liu and Yue Zhang. 2017. Attention modeling for targeted sentiment. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. 572–577.

[26] Yunfei Long, Qin Lu, Rong Xiang, Minglei Li, and Chu Ren Huang. 2017. A cognition based attention model for sentiment analysis. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

[27] Minh-Thang Luong, Hieu Pham, and Christopher D. Manning. 2015. Effective approaches to attention-based neural machine translation. arXiv:1508.04025.

[28] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*. 3111–3119.

[29] Saif M. Mohammad, Svetlana Kiritchenko, and Xiaodan Zhu. 2013. NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. In *Proceedings of the Joint Conference on Lexical and Computational Semantics*. 321–327.

[30] Preslav Nakov, Alan Ritter, Sara Rosenthal, Fabrizio Sebastiani, and Veselin Stoyanov. 2016. SemEval-2016 Task 4: Sentiment analysis in Twitter. In *Proceedings of the International Workshop on Semantic Evaluation*. 1–18.

[31] Sascha Narr, Michael Hulfenhaus, and Sahin Albayrak. 2012. Language-independent Twitter sentiment analysis. In *Proceedings of the Learning, Knowledge, and Adaption Conference (LWA'12)*. 12–14.

[32] Petra Kralj Novak, Jasmina Smailović, Borut Sluban, and Igor Mozetič. 2015. Sentiment of emojis. *PLoS One* 10, 12 (2015), e0144296.

[33] Debora Nozza, Elisabetta Fersini, and Enza Messina. 2017. A multi-view sentiment corpus. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*. 273–280.

[34] Alexander Pak and Patrick Paroubek. 2010. Twitter as a corpus for sentiment analysis and opinion mining. In *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC'10)*. 1320–1326.

[35] Yafeng Ren, Donghong Ji, and Han Ren. 2018. Context-augmented convolutional neural networks for Twitter sarcasm detection. *Neurocomputing* 308 (2018), 1–7.

[36] Sara Rosenthal, Preslav Nakov, Svetlana Kiritchenko, Saif Mohammad, Alan Ritter, and Veselin Stoyanov. 2015. Semeval-2015 Task 10: Sentiment analysis in Twitter. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval'15)*. 451–463.

[37] Jayashree Subramanian, Varun Sridharan, Kai Shu, and Huan Liu. 2019. Exploiting emojis for sarcasm detection. In *Proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction, and Behavior Representation in Modeling and Simulation*. 70–80.

[38] Ming Tan, Cicero dos Santos, Bing Xiang, and Bowen Zhou. 2016. Improved representation learning for question answer matching. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics: Volume 1, Long Papers*. 464–473.

[39] Duyu Tang, Bing Qin, and Ting Liu. 2015. Document modeling with gated recurrent neural network for sentiment classification. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. 1422–1432.

[40] Ye Tian, Thiago Galery, Giulio Dulcinati, Emilia Molimpakis, and Chao Sun. 2017. Facebook sentiment: Reactions and emojis. In *Proceedings of the 5th International Workshop on Natural Language Processing for Social Media*. 11–16.

[41] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems*. 5998–6008.

[42] Oriol Vinyals, Łukasz Kaiser, Terry Koo, Slav Petrov, Ilya Sutskever, and Geoffrey Hinton. 2015. Grammar as a foreign language. In *Advances in Neural Information Processing Systems*. 2773–2781.

[43] Joseph B. Walther and Kyle P. D́'Addario. 2001. The impacts of emoticons on message interpretation in computer-mediated communication. *Social Science Computer Review* 19, 3 (2001), 324–347.

[44] Xinyu Wang, Chunhong Zhang, Yang Ji, Li Sun, Leijia Wu, and Zhana Bao. 2013. A depression detection model based on sentiment analysis in micro-blog social network. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining*. 201–213.

[45] Alecia Wolf. 2000. Emotional expression online: Gender differences in emoticon use. *CyberPsychology & Behavior* 3, 5 (2000), 827–833.

[46] Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016. Hierarchical attention networks for document classification. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 1480–1489.

[47] Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2017. Hierarchical attention networks for document classification. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 1480–1489.

[48] Matthew D. Zeiler. 2012. ADADELTA: An adaptive learning rate method. arXiv:1212.5701.

[49] Lei Zhang, Shuai Wang, and Bing Liu. 2018. Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8, 4 (2018), e1253.

[50] Xinjie Zhou, Xiaojun Wan, and Jianguo Xiao. 2016. Attention-based LSTM network for cross-lingual sentiment classification. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. 247–256.

Received February 2019; revised December 2019; accepted March 2020