Postprint: Weibo Sentiment Analysis Model Based on Emoji Attention Mechanism

Authors: Tan Hao, Deng Shuwen, Qian Tao, Ji Donghong

Date: 2018-05-24T00:00:00+00:00

Abstract

To effectively identify sentiment polarity in Chinese microblogs, based on the cognitive fact that emoticons can alter or enhance the sentiment polarity of microblog text, we propose a neural network model for microblog sentiment analysis based on an emoticon attention mechanism. This model employs a Bidirectional Long Short-Term Memory (Bi-LSTM) network to learn feature representations of text, while utilizing an emoticon attention mechanism to obtain new feature representations that integrate text with emoticons, thereby achieving microblog sentiment recognition. Experimental results demonstrate that compared with a Bi-LSTM model that inputs both pure text and emoticons, the emoticon attention mechanism-based model improves accuracy by 4.06%; and compared with a Bi-LSTM model that only inputs pure text, the accuracy improvement reaches 6.35%.

Full Text

Emoji-Attentional Neural Network for Microblog Sentiment Analysis

Tan Hao[1], Deng Shuwen[2], Qian Tao[2], Ji Donghong[1] 1. School of Computer Science, Wuhan University, Wuhan 430072, China 2. School of Computer Science & Technology, Hubei University of Science & Technology, Xianning, Hubei 437100, China

Abstract: To effectively identify sentiment polarity in Chinese microblogs, this paper proposes an emoji-attentional neural network model for microblog sentiment analysis based on the cognitive fact that emojis can alter or intensify the sentiment polarity of textual content. The model employs a bidirectional recurrent neural network (Bi-LSTM) to learn feature representations of text, then utilizes an emoji attention mechanism to obtain new feature representations that integrate text with emojis, thereby achieving sentiment recognition for microblogs. Experimental results demonstrate that compared with a Bi-LSTM model that inputs plain text and emojis, the emoji-attentional model improves accuracy by 4.06%; compared with a Bi-LSTM model that only inputs plain text, the emoji-attentional model improves accuracy by 6.35%.

Keywords: emoji; microblog; sentiment analysis; attention mechanism

0 Introduction

Microblogs have become one of the most important media for people to exchange information and express emotions and opinions. Numerous research efforts have focused on microblog text for tasks such as public opinion analysis [1], emergency event detection [2], opinion mining [3], and sentiment analysis [4,5], with sentiment analysis being one of the foundational and key topics in microblog research. Current sentiment analysis primarily employs methods including Support Vector Machines [6], decision trees [7], and neural networks [8-10]. To express emotions and viewpoints, people use large numbers of emojis in microblogs, prompting many researchers to incorporate emojis as important factors in their models. Common approaches treat emojis as model features [11-12] or as a form of natural annotation [13-14] to expand training corpora. However, research [13] has shown that directly using emojis as sentiment labels introduces significant noise.

Studies in social and cognitive sciences suggest that emojis function similarly to non-verbal components in human communication, such as facial expressions and gestures, which convey emotions between communicators [15]. These studies argue that in social media, emojis exert the most significant influence on textual sentiment expression and can even alter the sentiment polarity of text. Consider the following two examples: 1. “Tomorrow I have to go to school” 2. “Mom’ s birthday gift for me”

The sentiment shift in these texts is illustrated in [Figure 1: see original paper]. As shown, emojis play a crucial role in textual sentiment expression. However, sentiment expression in microblogs is not simply the additive combination of text and emojis. For instance, although the two examples above contain the same emoji, its effect differs depending on the context. This effect resembles the attention mechanism in neural networks [16], where each word in a microblog sentence carries different information weights, and emojis simultaneously influence the weights of words in the text, transforming textual expression and consequently changing sentiment polarity.

Therefore, this paper combines these insights and proposes a neural network model based on emoji attention mechanisms. [Figure 2: see original paper] presents the framework of our proposed model. Microblogs are divided into two types of information: text and emojis. The model first employs a bidirectional recurrent neural network (Bi-LSTM) to learn low-dimensional feature representations of the input text, then combines these with emoji vectors through an emoji attention network layer to alter the weights of words in the text, thereby obtaining textual feature representations for final sentiment classification.

The main contributions of this paper are: (a) construction of an emoji-based sentiment annotation corpus for microblogs containing 6,905 posts; (b) unlike previous work that treated emojis as single features, this paper trains emojis and words simultaneously during corpus processing to obtain emoji word vectors containing contextual information, and proposes a neural network model based on emoji attention mechanisms. Experimental results demonstrate the effectiveness of this model.

1 Related Work

Sentiment analysis involves analyzing, processing, summarizing, and reasoning about subjective texts with emotional coloring. For example, analyzing movie reviews and product evaluations helps objectively assess products. It is typically treated as a classification problem, with various classification models widely applied, such as SVM, decision trees, and neural networks. Common features include words, word counts, part-of-speech tags, syntactic structures, and manually annotated sentiment dictionaries.

In recent years, neural network models have developed rapidly, avoiding complex feature engineering and achieving widespread application in sentiment analysis. Most methods employ RNN, CNN, and other models. Chen et al. [16] proposed a user-product attention mechanism for sentiment analysis tasks. Zhang et al. [17] introduced a gating mechanism. These models primarily use text as input without considering the effect of emojis on text.

In social media platforms such as Sina Weibo and Twitter, numerous emojis exist to express user emotions and opinions. Current strategies for utilizing emojis mainly include three categories:

The first strategy [13,14] treats emojis as natural annotations, assuming that emojis independently express user emotions and opinions. Purver et al. [18] used hashtags and emotions in Twitter data to generate training datasets. Experimental results showed this method effective when using happiness, sadness, and anger as emotion labels, but less effective for other emoji types. Davidov et al.' s research [13] revealed that due to emoji ambiguity—where the same emoji can be positive or negative—training corpora built using emojis contain considerable noise that may negatively impact model training.

The second strategy [11,12] incorporates emojis as features into analysis models. Jiang et al. [19] fused microblog structural features, sentence structure features, and emoji features in an SVM model for three-class sentiment classification (positive, neutral, negative). This strategy also fails to capture the emotional effect of emojis on text.

The third strategy treats emojis and text as two parallel information sources. Hogenboom et al. [20] divided social media text into emoji and text components, then used different models to compute sentiment for each, finally linearly combining the two sentiments to obtain the final text sentiment.

Although these three strategies consider emojis in models, they do not account for the mechanism by which emojis affect text. Attention mechanisms were initially applied in image processing [21] and have recently been gradually introduced to natural language processing, serving to select key information from numerous sources. In natural language processing, research has shown that attention mechanisms significantly improve performance in machine translation [22], question answering systems [23], and sentiment analysis [16]. This paper proposes a neural network model based on emoji attention mechanisms, effectively simulating the cognitive fact that emojis influence textual sentiment.

2 Model

2.1 Bi-LSTM-Based Sentiment Classification Model

The Bidirectional Long Short-Term Memory (Bi-LSTM) model is a special recurrent neural network (RNN) primarily designed to process sequential data and has achieved widespread application in natural language processing. In sentiment classification tasks, this model is typically used to learn sentence representations for subsequent text classification. Yang et al. [24] used Bi-LSTM for document classification with good results. Below is a brief introduction to the Bi-LSTM model.

LSTM is a special recurrent neural network that enhances the memory mechanism for receiving input information and training data, substantially improving output results. For the t-th word w_t in short text, the model first maps w_t to a word vector $x_t \in \mathbb{R}^d$, where d is the word vector dimension. The input to an LSTM unit is: word x_t, previous hidden state h_{t-1}, and previous memory state c_{t-1}; the output is h_t, c_t. The specific formulas are as follows:

$$\begin{aligned}
i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i) \\
f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f) \\
\tilde{c}_t &= \tanh(W_c x_t + U_c h_{t-1} + b_c) \\
c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\
o_t &= \tanh(W_o x_t + U_o h_{t-1} + b_o) \\
h_t &= o_t \odot \tanh(c_t)
\end{aligned}$$

For standard LSTM, a disadvantage exists: it can only read text forward. Therefore, this paper uses the Bi-LSTM model that can read text bidirectionally. Bi-LSTM contains a forward $\overrightarrow{LSTM}$ that reads from x_1 to x_T, and a backward $\overleftarrow{LSTM}$ that reads from x_T to x_1:

$$\overrightarrow{h_t} = \overrightarrow{LSTM}(x_t, \overrightarrow{h_{t-1}}), \quad t \in [1, T]$$

$$\overleftarrow{h_t} = \overleftarrow{LSTM}(x_t, \overleftarrow{h_{t+1}}), \quad t \in [T, 1]$$

After this layer, the model transforms a word w_t into a hidden variable h_t, where $h_t = \overrightarrow{h_t} \oplus \overleftarrow{h_t}$. Concatenating h_t yields the representation of a single word, i.e., $s = [h_1, h_2, ..., h_T]$. This is then passed through an average pooling layer to obtain the representation of a single sentence.

To perform sentiment classification on text, the model uses a nonlinear function to map the text representation to a target category space. After introducing the emoji attention mechanism, the model obtains the final vector representation of the text, then employs a softmax function to obtain its sentiment distribution.

2.2 Emoji-Attention-Based Bi-LSTM Sentiment Classification Model

To represent the mechanism by which emojis affect text, this paper proposes an emoji attention mechanism to model emoji effects on text. For a microblog post, each word has a different impact on sentiment polarity, and the effect of interacting with emojis also differs. The emoji attention mechanism measures the importance weights of words in microblogs after combining them with emojis.

For a microblog text $\{w_1, w_2, ..., w_T; E_1, E_2, ..., E_k\}$, where w_i represents tokenized words in the microblog and E_j represents emojis in the microblog. First, both w_i and E_j are converted to vector representations. Let $x_t \in \mathbb{R}^d$ be the word vector representation and $e_j \in \mathbb{R}^d$ be the emoji vector representation, where d is the vector dimension.

Since many microblog users post multiple identical emojis in the same post, the average of emoji vectors in the microblog is taken here to prevent excessive weight of a single emoji from affecting model performance:

$$\bar{e} = \frac{1}{k} \sum_{j=1}^{k} e_j$$

Let h_t denote the representation of pure text w in the microblog after passing through the Bi-LSTM neural network, as described in Section 2.1. The representation of the entire sentence is as follows:

$$s = \sum_{t=1}^{T} \alpha_t h_t$$

Here, α_t represents the importance degree of the t-th word combined with emojis, i.e., the attention weight of the t-th word in the sentence as modeled by the attention mechanism. α_t is defined as:

$$\alpha_t = \frac{\exp(\mathrm{score}(\bar{e}, h_t))}{\sum_{j=1}^{T} \exp(\mathrm{score}(\bar{e}, h_j))}$$

The score function measures the importance of word-emoji combinations and is defined as:

$$\mathrm{score}(\bar{e}, h_t) = \tanh(W_H h_t + W_E \bar{e} + b)^\top v$$

where W_H, W_E are weight matrices, b is a bias vector, v is a weight vector, and $\top$ denotes transposition.

2.3 Optimization Objective

The model adopts cross-entropy loss as its loss function. If D represents the training microblog set, the loss function for the emoji-attention-based model is expressed as:

$$L = -\sum_{c=1}^{C} \sum_{d \in D} \log(p_d(c))$$

where C represents the number of sentiment labels and $p_d(c)$ represents the predicted probability of sentiment label c.

3 Experiments

3.1 Dataset

To obtain training and testing corpora, 100,000 microblogs containing emojis were extracted from Sina Weibo in April 2017. The jieba Chinese word segmenter was used for tokenization, replacing URLs, usernames, and topic hashtags in microblogs. Posts with length less than 5 were filtered out. Then over 10,000 microblog texts were randomly sampled as candidates for annotation, with the requirement that each emoji appear more than 10 times. After removing duplicates and posts with too few emojis, 6,905 posts remained in the corpus. During annotation, sentiment labels were divided into three categories: positive, neutral, and negative. The corpus and annotation statistics are shown in .

Table 1: Corpus Statistics | Different Emoji Count | Total Emoji Occurrences | Number of Instances | |—————-|——————|—————|

To obtain vector representations of words and emojis in microblogs, 100M microblog posts were extracted for pre-training. After tokenization and preprocessing as described in Section 3.1, ordinary words and emoji symbols were treated as the same type of token. Word vector representations were pre-trained using the SkipGram model [25] with a dimension of 300, yielding vector representations for microblog words and emoji vectors containing contextual information. For words or emojis not appearing in the vocabulary during model operation, random initialization was used.

Microblog text representations were learned using Bi-LSTM. Each hidden node' s vector was 100-dimensional, with each word' s representation formed by concatenating bidirectional hidden vectors, resulting in a 200-dimensional output representation per word. The Adadelta method [26] was used to optimize parameters during training. In each optimization step, the model updated weight matrices, bias vectors, word vectors, and emoji vectors. Training ran for 100 epochs.

The annotated corpus was split into training, development, and test sets in an 8:1:1 ratio. Accuracy was used as the evaluation metric.

3.3 Experimental Results

The primary goal of these experiments was to test the effect of emojis on text. Therefore, three baselines were established:

a) **Emoji-based**: Using only emoji polarity for judgment. Each emoji' s polarity was manually annotated.

b) **Bi-LSTM-text**: Pure text input to Bi-LSTM sentiment analysis network. Emojis were removed from microblogs, with pure text as network input. The microblog representation from Bi-LSTM served directly as the feature vector for the classification model. Optimization objectives and training methods were identical to our proposed model.

c) **Bi-LSTM-text-emoji**: Pure text + emoji input to Bi-LSTM sentiment analysis network. Emojis were retained in microblogs, with both emojis and text as network input. The microblog representation from Bi-LSTM served directly as the feature vector for the classification model. Optimization objectives and training methods were identical to our proposed model.

Table 2: Sentiment Recognition Results

Model	Accuracy (%)
Emoji polarity only	
Bi-LSTM-text	68.51
Bi-LSTM-text-emoji	70.80
Emoji-Attention Bi-LSTM	74.86

Table 2 shows the recognition accuracy of each model. The results reveal that introducing emojis into the model improves performance from 68.51% (text-only) to 70.80%. This is primarily because microblog text contains many humorous, self-deprecating, and sarcastic posts where emojis cause shifts in sentiment polarity that cannot be identified from text alone. Comparing attention-based and non-attention emoji models, the attention-based model accuracy rises to 74.86%, demonstrating that emoji attention mechanisms more effectively represent text sentiment.

Table 3: Recognition Results for Different Sentiment Polarity Texts Using Emoji-Attention Model | Polarity | Precision | Recall | F1-Score | | ——-|——-|—-|——-| | Positive | | | | | Neutral | | | | | Negative | | | |

Table 3 shows that the emoji-attention mechanism model outperforms the emoji-as-feature model. To verify the attention model’ s impact on microblogs with different sentiment polarities, Table 3 presents recognition precision for different sentiment polarities. The results show that after adding emojis, negative polarity accuracy improves significantly, also exceeding positive polarity recognition rates, while neutral polarity accuracy remains lowest. This indicates negative emojis have the greatest impact on text polarity; emojis have little effect on neutral microblog accuracy but significantly improve recall.

The following examples show cases where the Emoji-Attention Bi-LSTM (LSTM-EA) model correctly identified sentiment while Bi-LSTM-text+emoji (LSTM-E) failed:

1) “I don’ t mind PDA, but I mind PDA flooding my feed. Sorry, us single dogs are just that direct, okay?”
2) “Sorry, my comprehension is low. Your wake-up call didn’ t wake me up; instead, it knocked away my respect and gratitude for you. Thanks.”
3) “[Video of two cars colliding then drivers laughing it off] PS: Fellow Zhumadian folks, can’ t understand~ Hahaha∼Reminds me of last time when heavy rain in Nanping trapped me for 3 days and I fished from my balcony, my fellow Chongqing people are so cute. Also 鄙视 those who want to capitalize on tragedy and spread rumors.”
4) “Thanks for passengers’ support and understanding!”

Examples 1) and 2) are clear cases of sarcasm where the LSTM-E model, which simply averages word vectors as classifier input, makes incorrect predictions. The LSTM-EA model correctly predicts by considering different weights for each word. Examples 3) and 4) are misclassified by LSTM-E because they only consider some negative words while ignoring the overall sentiment requiring holistic interaction.

Since microblogs contain much sarcasm and “twist-ending” examples like those above, the proposed model demonstrates superiority and practical value.

3.4 Experimental Analysis

The experiments above demonstrate that our proposed model significantly outperforms models considering only pure text and also surpasses models that simply treat emojis as features. The main reasons are that emojis change or intensify text polarity, while the emoji attention model also discovers semantic weights within the text, improving recognition effectiveness.

In Table 2, neutral sentiment microblogs still show lower accuracy because they contain some biased words and emojis while the text itself is not 倾向性; another reason is that some users employ emojis casually, using emojis unrelated to the main text. The following two sentences are actually neutral but were identified as positive:

"Life is like a volcano. When you' re angry and furious, it erupts; when you' re happy and joyful, it sprays underground springs to water the flowers and plants at your feet."

"Can' t be careless···"

4 Conclusion

Based on the cognitive fact that emojis play an important role in textual sentiment expression, this paper proposes a sentiment analysis model based on emoji attention mechanisms. In the model, each word in a sentence receives different weights after combining with emojis, enabling more precise semantic expression of a sentence. Experimental results show that the emoji attention mechanism can effectively identify microblog sentiment polarity, with both accuracy and F-score surpassing methods that simply add emojis as features to neural networks, achieving good results.

This work still has some limitations. During annotation, microblog polarity was labeled with only three categories, which is relatively coarse-grained. Additionally, some users' microblog text and emoji usage are rather casual, affecting model performance. Future work will investigate fine-grained microblog sentiment analysis based on emojis.

References

[1] Kannangara S. Mining Twitter for fine-grained political opinion polarity classification, ideology detection and sarcasm detection [C]// Proc of the 11th ACM International Conference on Web Search and Data Mining. New York: ACM Press, 2018: 751-752.

[2] Shi Si, Jin Dawei, Goh T T. Real-time public mood tracking of Chinese microblog streams with complex event processing [J]. IEEE Access, 2017, 5: 421-431.

[3] Yin Fulian, Zhang Beibei, Huang Bochen, et al. An public opinion trend prediction method based on neural network algorithm [C]// Proc of the 2nd IEEE International Conference on Computational Intelligence and Applications. Piscataway, NJ: IEEE Press, 2017: 440-444.

[4] Peng Haiyun, Cambria E, Hussain A. A review of sentiment analysis research in Chinese language [J]. Cognitive Computation, 2017, 9 (4): 423-435.

[5] Cortis K, Freitas A, Daudert T, et al. Semeval-2017 task 5: fine-grained sentiment analysis on financial microblogs and news [C]// Proc of the 11th International Workshop on Semantic Evaluation. Stroudsburg, PA: ACL, 2017: 519-535.

[6] Pang Bo, Lee L, Vaithyanathan S. Thumbs up?: Sentiment classification using machine learning techniques [C]// Proc of ACL Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2002: 76-86.

[7] Prasad S S, Kumar J, Prabhakar D K, et al. Sentiment classification: an approach for indian language tweets using decision tree [C]// Proc of International Conference on Mining Intelligence and Knowledge Exploration. Cham: Springer, 2015: 656-663.

[8] Tang Duyu, Qin Bing, Liu Ting. Document modeling with gated recurrent neural network for sentiment classification [C]// Proc of Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2015: 1422-1432.

[9] Dos Santos C N, Gatti M. Deep convolutional neural networks for sentiment analysis of short texts [C]// Proc of the 25th International Conference on Computational Linguistics: Technical Papers. Stroudsburg, PA: ACL, 2014: [pages].

[10] Tai Kaisheng, Socher R, Manning C D. Improved semantic representations from tree-structured long short-term memory networks [EB/OL]. (2015) [2018-04-10]. https://arxiv.org/pdf/1503.00075.

[11] Zhao Jichang, Dong Li, Wu Junjie, et al. Moodlens: an emoticon-based sentiment analysis system for chinese tweets [C]// Proc of the 18th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2012: 1528-1531.

[12] Ptáček T, Habernal I, Hong Jun. Sarcasm detection on czech and english twitter [C]// Proc of the 25th International Conference on Computational Linguistics: Technical Papers. Stroudsburg, PA: ACL, 2014: 213-223.

[13] Davidov D, Tsur O, Rappoport A. Enhanced sentiment learning using twitter hashtags and smileys [C]// Proc of the 23rd International Conference on Computational Linguistics: Posters. Stroudsburg, PA: ACL, 2010: 241-249.

[14] Go A, Bhayani R, Huang Lei. Twitter sentiment classification using distant supervision, CS224N [R]. Stanford, 2009.

[15] Walther J B, D' Addario K P. The impacts of emoticons on message interpretation in computer-mediated communication [J]. Social Science Computer Review, 2001, 19 (3): 324-347.

[16] Chen Huiming, Sun Maosong, Tu Cunchao, et al. Neural sentiment classification with user and product attention [C]// Proc of Conference on Empirical Methods in Natural Language Processing. Stroudsburg, PA: ACL, 2016: 1650-1659.

[17] Zhang Meishan, Zhang Yue, Vo D T. Gated neural networks for targeted sentiment analysis [C]// Menlo Park, CA: AAAI Press, 2016: 3087-3093.

[18] Purver M, Battersby S. Experimenting with distant supervision for emotion classification [C]// Proc of the 13th Conference of the European Chapter of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2012: [pages].

[19] Jiang Fei, Cui Anqi, Liu Yiqun, et al. Every Term Has Sentiment: Learning from emoticon evidences for Chinese microblog sentiment analysis [C]// Natural Language Processing and Chinese Computing. Berlin: Springer, 2013: 224–235.

[20] Hogenboom A, Bal D, Frasincar F, et al. Exploiting emoticons in sentiment analysis [C]// Proc of the 28th Annual ACM Symposium on Applied Computing. New York: ACM Press, 2013: 703-710.

[21] Mnih V, Heess N, Graves A. Recurrent models of visual attention [C]// Advances in Neural Information Processing Systems. 2014: 2204–2212.

[22] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate [EB/OL]. (2016) [2018-04-10]. https://arxiv.org/pdf/1409.0473.

[23] Hermann K M, Kocisky T, Grefenstette E, et al. Teaching machines to read and comprehend [C]// Advances in Neural Information Processing Systems. 2015: 1693-1701.

[24] Yang Zichao, Yang Diyi, Dyer C, et al. Hierarchical attention networks for document classification [C]// Proc of Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg, PA: ACL, 2016: 1480-1489.

[25] Mikolov T, Sutskever I, Chen Kai, et al. Distributed representations of words and phrases and their compositionality [C]// Advances in Neural Information Processing Systems. 2013: 3111-3119.

[26] Zeiler M D. Adadelta: an adaptive learning rate method [EB/OL]. (2012-12-22) [2018-04-10]. https://arxiv.org/pdf/1212.5701.

Note: Figure translations are in progress. See original paper for figures.