Postprint: Research on Feature Engineering for Uyghur Sentiment Classification

Authors: Tuerhongtai Rexidanmu, Wushour Silamu

Date: 2018-10-11T00:00:00+00:00

Abstract

Due to the current lack of systematic research on feature representation for Uyghur sentiment classification, new features and their combinations were extracted from a Uyghur sentiment-annotated corpus at different scales based on traditional n-gram features, and positive/negative sentiment classification was performed on the Uyghur sentiment corpus using a Support Vector Machine (SVM) classifier. Experimental results demonstrate that among the extracted basic features, unigram features achieve the optimal classification efficiency; the combination of unigram features and phrase features can further improve classification efficiency, with the best classification performance showing a 1.78% improvement over that of unigram features. For the first time, a comprehensive evaluation of the classification performance of different features was conducted on a unified annotated dataset, and the research findings can provide guidance for future Uyghur sentiment classification studies.

Full Text

Preamble

Research on Feature Construction for Uyghur Text Sentiment Classification

Rexidanmu Tuerhongtai[1],[2], *Wushour Silamu*[1]

[1]College of Information Science & Engineering, Xinjiang University, Urumqi 830046, China
[2]College of Electronic & Information Engineering, Yili Normal University, Yili, Xinjiang 835000, China

Abstract: Due to the lack of systematic research on feature representation for Uyghur sentiment classification, this study extracts new features and their combinations from a Uyghur sentiment-annotated corpus at different scales, building upon traditional n-gram features. Using a Support Vector Machine (SVM) classifier, we perform binary sentiment classification (positive/negative) on the Uyghur sentiment corpus. Experimental results demonstrate that among the extracted basic features, unigram features achieve the best classification performance. The combination of unigram features with phrase features can further improve classification efficiency, with the optimal combination achieving a 1.78% improvement over unigram features alone. This work represents the first comprehensive evaluation of different feature types on a unified annotated dataset, providing guidance for future Uyghur sentiment classification research.

Keywords: sentiment classification; feature construction; combined features; Uyghur

0 Introduction

Text sentiment classification is fundamentally a text classification problem [1]. Since Pang et al. [2] from Cornell University first applied machine learning techniques to sentiment classification, this approach has gained widespread attention and rapid development. Machine learning-based classification methods have undergone two major waves: shallow learning (traditional learning) [3,4] and deep learning [5,6].

In recent years, deep learning methods have been increasingly applied to text sentiment classification research. Deep learning models automatically learn word vectors from large-scale unannotated corpora and use them as basic features, overcoming the limitations of traditional methods that rely on manually designed features and reducing human labor and time costs. However, a critical issue arises when using word vectors trained by deep learning models as input features for sentiment classification: since word vectors are constructed based on lexical context without considering sentiment information, words with similar contexts but opposite sentiment polarities may learn similar vector representations, potentially reducing both the efficiency and quality of sentiment classification. To address this problem, researchers have combined sentiment word vectors with traditional manually designed features to improve deep learning model performance.

Traditional machine learning-based sentiment classification research has achieved fruitful results. Following the feature engineering paradigm, numerous studies have conducted in-depth investigations into feature representation for sentiment classification, systematically examining various feature types such as unigram, bigram, and other common bag-of-words features, as well as syntactic, semantic, negation, and combined features.

Compared to sentiment classification research in languages like Chinese and English, Uyghur text sentiment classification remains in its infancy. Uyghur is a morphologically rich agglutinative language with far more complex morphological structures than Chinese or English. Consequently, sentiment classification of Uyghur text must consider not only general technical issues but also the unique characteristics of the Uyghur language itself. Currently, Uyghur text sentiment classification research is in an exploratory stage, lacking systematic studies on feature representation. Therefore, most research in this area must start from scratch.

This paper extracts unigram, bigram, trigram, sentiment words, part-of-speech features, bi-tagged features, and generalized bi-tagged features from our self-constructed Uyghur sentiment-annotated corpus [7]. Using Mutual Information (MI) for feature selection, we extract optimal features and form combined features including unigram+bigram, unigram+bi-tagged, and unigram+generalized bi-tagged. We then evaluate and compare the performance of different features for Uyghur sentiment classification on this sentiment-annotated dataset.

This work extracts both traditional n-gram features and multi-word features that capture semantic relationships between words, and for the first time systematically evaluates the performance of different features on a unified annotated dataset. The findings can provide guidance for future Uyghur text sentiment classification research and serve as a reference for related languages such as Kazakh and Kyrgyz.

---

1 Related Work

Traditional machine learning-based sentiment classification methods use manually annotated sentiment texts as training data, extract sentiment features, construct sentiment classifiers using machine learning techniques, and then apply the trained classifiers to new documents. The classification efficiency of these methods heavily depends on the quality of sentiment features. Numerous studies have systematically investigated how different features affect sentiment annotation.

Habernal et al. [3] conducted sentiment classification experiments on their self-constructed Czech Social Media sentiment corpus and movie/product review corpora, verifying the impact of several preprocessing methods on classification efficiency. They extracted n-gram, character n-gram, part-of-speech, and emoticon features, as well as combination features derived from them, and performed sentiment classification using Support Vector Machine (SVM) and Maximum Entropy (MaxEnt) classifiers. Rehab et al. [4] studied the effects of stemming, feature combination, and n-gram models on classification results, applying SVM, Naive Bayes (NB), and K-Nearest Neighbors (KNN) classifiers on two datasets, all achieving favorable results.

In Chinese sentiment classification research, Li Zekui et al. [8] conducted comparative studies on word, phrase, numerical, and syntactic features based on

Chinese microblog corpora, proposing a new feature based on dictionary rules for sentiment scoring, and through extensive experiments and analysis, identified reliable feature combinations.

With the successful application of deep learning methods in image processing and speech recognition, an increasing number of researchers have recently applied these methods to sentiment classification tasks. Kim [5] used Convolutional Neural Networks (CNN) for sentiment analysis and question classification, achieving good results. Liang Jun et al. [6] utilized recursive autoencoder models to automatically learn task-relevant features, avoiding manual feature selection, and demonstrated through comparative experiments that this model could improve sentiment classification accuracy.

Some researchers, considering the lack of sentiment information in deep learning-based word vector features, have combined traditional manually designed features with deep learning features to improve classification efficiency. Sun Chaohong [9] fused shallow features such as unigram, part-of-speech (POS), and sentiment dictionary features with word vectors trained by Word2vec, using an LSTM-improved RNN model for sentiment polarity classification of microblog texts. Xu Yingying [10] combined word vectors with traditional manual features to construct a supervised ranking model for predicting sentiment intensity, which won first place in the English phrase sentiment intensity prediction task at the 2016 SemEval (International Workshop on Semantic Evaluation) competition.

In Uyghur text sentiment classification, Tian Shengwei et al. [11] selected unigram, bigram, and trigram features, employed feature selection methods such as document frequency, chi-square test, and information gain, and conducted research using Naive Bayes, SVM, and Maximum Entropy classifiers. Reyilaimu Paerhati et al. [12] performed binary classification on a small self-constructed corpus by extracting discriminative words. Abudusalamu Dawuti et al. [13] combined discriminative words extracted from [12] with a sentiment dictionary for sentiment classification, achieving better results. Li Min et al. [14] studied Uyghur text sentiment classification based on stacked autoencoder neural networks, obtaining higher accuracy than traditional machine learning algorithms, with a macro-accuracy of 90.5%. Li Dongbai et al. [15] obtained vector representations for each word in the corpus through word2vec, linearly combined word vectors with part-of-speech features, used stacked autoencoder algorithms to automatically learn features from large-scale unannotated implicit sentiment texts, and completed automatic classification of implicit sentiment in Uyghur texts through a softmax classifier. Wang Shuheng et al. [16] implemented a Uyghur sentiment classification model based on word embedding and bidirectional LSTM, combining Uyghur linguistic features and sentiment features between words, with experimental results outperforming RNN, CNN, and SVM classifiers.

2 Experimental Dataset and Preprocessing

2.1 Uyghur Review Corpus

The corpus [7] consists of user comments on various topics collected from several major Uyghur websites. Since these comments contain rich sentiment information, they satisfy the requirements for a sentiment corpus. Each comment in the corpus is annotated with one of three sentiment orientations: positive, negative, or neutral. The corpus contains a total of 15,814 annotated comments, including 10,368 positive, 4,515 negative, and 931 neutral comments. Detailed information about the corpus is shown in Table 1 .

Since this study focuses only on binary sentiment classification (positive and negative), we selected 4,515 positive comments and 4,515 negative comments from the annotated corpus as our experimental data.

2.2 Corpus Preprocessing

Uyghur exhibits very rich morphological variations and a vast vocabulary. Although the number of stems and affixes in Uyghur is limited, they can theoretically combine to form an infinite number of words, with the vast majority of words appearing only once in the corpus [17]. This leads to extremely high feature space dimensionality and severe data sparsity problems in Uyghur natural language processing tasks. Therefore, preprocessing of the experimental data is necessary.

2.2.1 Uyghur Morphological Analyzer The Uyghur morphological analyzer is a preprocessing tool developed by the Multilingual Information Processing Laboratory of Xinjiang University. It implements multiple functions including sentence boundary identification, stemming, and part-of-speech tagging. The tool uses a combination of statistical and rule-based methods to identify sentence boundaries [18]. In stemming, each word is represented as a tree structure, with root nodes representing stems, child nodes representing affixes, and edges representing constraints between stems and affixes. The stemming process fully considers phonological changes that occur during morphological variations in Uyghur [19]. Part-of-speech tagging implements 15 first-level annotation specifications as shown in Table 2 .

A Uyghur sentence processed by the morphological analyzer yields results where "T=" indicates part-of-speech, "S=" indicates the word root, and marks the end of a sentence.

2.2.2 Preprocessing Steps **1) Word Segmentation:** To obtain word features, text must first be segmented. Uyghur is an alphabetic language where words are separated by spaces and punctuation marks. Therefore, word segmentation for Uyghur is not a technical challenge and can be accomplished using spaces and punctuation marks.

2) Stemming: In Uyghur, the stem is the main component expressing lexical semantics, while morphological suffixes express grammatical and tense information [21]. To reduce feature space dimensionality and avoid the curse of dimensionality, stemming is required. After stemming, the basic semantics of the original word are preserved while effectively reducing feature space dimensionality. For example, the word “ ” (school) can form different words with slightly different spellings but the same core meaning by connecting various affixes (e.g., , , , , .(We applied the Uyghur morphological analyzer to perform stemming on our corpus.

3) Part-of-Speech Tagging: POS information is considered an important clue for sentiment expression. Adjectives, adverbs, verbs, and nouns can carry important sentiment information. Our experiments extract words of different POS categories as features. Additionally, since our proposed bi-tagged features are extracted based on POS combination rules in text, we used the Uyghur morphological analyzer for POS tagging on our corpus.

4) Stopword Removal: Uyghur sentiment texts contain many high-frequency words that do not contribute to sentiment classification, such as “man” (I), “u” (he/she), “bir” (one), “silar” (you plural). Using these words as text features would increase feature space dimensionality and degrade classifier performance. Therefore, it is necessary to remove these words. We eliminated non-sentiment-bearing stopwords from the text using a self-constructed Uyghur sentiment classification stopword list containing 1,305 words.

3 Feature Construction

3.1 Basic Features

1) n-gram Features: We extracted unigram, bigram, and trigram features from sentences, denoted as F_{uni}, F_{bi}, and F_{tri} respectively.

2) Sentiment Lexicon Features: Sentiment words typically contain rich emotional color and often reveal the attitude and emotion expressed in text, making them important features. We used all positive and negative words from our self-constructed Uyghur sentiment dictionary [22] as basic sentiment features, denoted as F_{dict}.

3) POS Features: POS information has long been considered an important indicator for sentiment expression. Therefore, we selected nouns, verbs, adjectives, adverbs, and interjections as basic sentiment features, denoted as F_{pos}.

4) bi-tagged Features: Traditional n-gram feature extraction methods generate high-dimensional feature vectors, which not only increase classification difficulty but also prolong classification time. Inspired by [23], we summarized several POS combination rules and extracted phrases consisting of two adjacent words with specific sequential relationships from the text, naming them bi-tagged features and denoting them as F_{bitag}. The POS combination rules for bi-tagged features are shown in Table 3 .

5) Generalized bi-tagged Features: While bigram, trigram, and bi-tagged features can extract contextual semantic information, they suffer from excessive sparsity, weak generalization ability, and high dimensionality. For example, consider these two sentences from training data: "Bügün shundaq tesirlik kënodin birni kurdüm, silerningmu kurüp bëqishinglarni tewsiye qiliman." (I watched a very touching movie today, and I recommend you watch it too.)
"Silerge bir tesirlik kitap tonöshtöray." (Let me introduce you to a touching book.)

From the training corpus, we can extract bigram and bi-tagged features such as "tesirlik këno" (touching movie), "shundaq tesirlik" (very touching), and "tesirlik kitap" (touching book). While these features effectively express sentiment, if the feature "tesirlik hë ekay" (touching story) appears in the test corpus, a classifier trained on the above training data cannot determine its sentiment orientation. To address this issue, following the approach in [24,25], we replace one of the two words in our extracted bi-tagged features with its corresponding POS tag, naming these generalized bi-tagged features and denoting them as F_{Gbitag}. In the above example, if we replace the second word in features like "tesirlik këno" , "tesirlik hë ekay" , and "tesirlik kitap" with their POS tags, we obtain the feature "tesirlik N" (touching N), achieving generalization across different features and effectively ensuring the generalization performance of most features in the training corpus. We experimented with both prefix replacement (F^h_{Gbitag}, replacing the first word in the pair with its POS tag) and suffix replacement (F^t_{Gbitag}, replacing the second word with its POS tag), determining the optimal replacement method through comparative experiments.

3.2 Combined Features

While unigram features consistently outperform other features in sentiment classification tasks, their limitation lies in the inability to extract contextual information from text. Although phrase features like bigram, trigram, and bi-tagged can enhance semantic content, they reduce the statistical quality of feature vectors, making features sparser and making it difficult for machine learning algorithms to extract useful statistical patterns for classification. Due to these drawbacks, the sentiment classification performance using these features is inferior to that of unigram features [3].

To address this issue, we combined F_{uni} with F_{bi} features, F_{bitag} features, and F_{Gbitag} features, forming F_{uni-bi}, $F_{uni-bitag}$, and $F_{uni-Gbitag}$ combined features respectively.

When combining features, we designed a combination ratio control parameter α, which determines the proportion of each feature type in the combined feature set. α represents the proportion of phrase features (combined with unigram features) in the total combined features, thereby determining the importance of each feature type in the combination. Taking $F_{uni-bitag}$ as an example, when the total number of combined features is N, the number of bi-tagged features is $N_{bi-tag} = N \times \alpha$, and the number of unigram features is $N_{uni} = N - N_{bi-tag}$, meaning we use N_{uni} unigram features and N_{bi-tag} bi-tagged features for combination.

---

4 Experiments and Results Analysis

After extracting different feature types from the Uyghur sentiment corpus, we used Mutual Information (MI) for feature selection and tf-idf for feature weighting. We then evaluated and compared the discriminative ability of different features using an SVM machine learning classifier for binary sentiment classification (positive/negative) on the Uyghur review sentiment corpus. Experiments employed 10-fold cross-validation: the dataset was divided into 10 subsets, with one subset used as the test set and the remaining nine as the training set in each round, for a total of 10 rounds. The final results were averaged across all rounds. All experiments were conducted using Python and the Scikit-learn toolkit, with classification performance evaluated using Accuracy.

4.1 Results on Basic Features

To verify the performance of our extracted features in Uyghur sentiment classification, we extracted different features from the corpus, used MI for feature ranking, and sequentially selected features ranked from top 10% to 90% to compare the impact of different feature set sizes on classifier performance. The experimental results are described in Table 4 .

As shown in Table 4, F_{uni} features achieve the best classification performance among all basic features. When selecting the top 30% of features (3,324 features), the classifier reaches its peak accuracy of 89.47%, but accuracy declines from this peak as the number of features increases. In our experiments, F_{bi} and F_{tri} features also achieved relatively good results. For example, when extracting 90% of features from F_{bi} (2,316 features), classification accuracy reaches 86.50%; when extracting 50% of features from F_{tri} (5,269 features), accuracy reaches 85.78%.

Among basic features, F_{dict} and F_{pos} features performed below expectations. Among phrase features, our proposed F_{bitag} features outperform F_{bi} features, achieving a maximum classification accuracy of 83.99%, which is 4.82% higher than F_{bi}' s maximum accuracy of 79.17%. Generalizing bi-tagged features can further enhance their classification performance. Both types of F_{Gbitag} features outperform F_{bitag}. For example, when extracting the top 60% of features from F^{t}_{Gbitag}, classification accuracy reaches 85.23%, representing a 1.24% improvement over F_{bitag} with the same number of features.

4.2 Results on Combined Features

To verify the performance of our combined features in Uyghur sentiment classification, we conducted sentiment classification experiments on the Uyghur review sentiment corpus using three types of combined features (F_{uni-bi}, $F_{uni-bitag}$, and $F_{uni-Gbitag}$). Considering that $F^t_{uni-Gbitag}$ outperformed $F^h_{uni-Gbitag}$ in basic feature experiments, we used $F^t_{uni-Gbitag}$ to form the combined feature $F_{uni-Gbitag}$. During experiments, we gradually increased the total number of features from 1,000 to 10,000 in increments of 1,000, while increasing the feature ratio control parameter α from 10% to 90%. The classification accuracies of the three combined features on the corpus are shown in Tables 5 through 7 . Due to space limitations, we only present results at selected intervals for α.

In experiments based on combined feature F_{uni-bi} (Table 5), when the total number of features is 8,000 and F_{bi} features account for 50% of the total, the classifier achieves its highest accuracy of 90.72%, representing a 1.25% improvement over F_{uni}' s maximum accuracy.

In experiments based on combined feature $F_{uni-bitag}$ (Table 6), when the total number of features is 5,000 and F_{bitag} features account for 50% of the total, the classifier achieves its highest accuracy of 90.97%, representing a 1.23% improvement over F_{uni}' s maximum accuracy and a 0.25% improvement over F_{uni-bi}' s maximum accuracy.

In experiments based on combined feature $F_{uni-Gbitag}$ (Table 7), when the total number of features is 8,000 and F_{Gbitag} features account for 50% of the total, the classifier achieves its highest accuracy of 91.25%, representing improvements of 1.78%, 0.53%, and 0.28% over F_{uni}, F_{uni-bi}, and $F_{uni-bitag}$ features respectively.

Experimental results demonstrate that combining unigram features with phrase features containing contextual semantic information can effectively overcome the individual limitations of each feature type and achieve better classification results than using any single feature alone. In our combined feature experiments, classification performance is optimal when unigram features and phrase features each account for 50% of the total features. This is because when unigram features suffer from data sparsity, phrase features can extract sentiment-rich contextual information that complements the unigram features.

Among the three combined features, $F_{uni-Gbitag}$ achieves the best classification performance. The primary reason is that bi-tagged features can eliminate many noisy features present in bigram features while extracting structurally stable and semantically complete contextual information. Generalizing bi-tagged features can further improve their statistical properties and effectively address data sparsity issues, which explains why $F_{uni-Gbitag}$ outperforms $F_{uni-bitag}$.

5 Conclusion

Addressing the lack of systematic research on feature representation for Uyghur text sentiment classification, this paper extracts eight types of basic features and three types of combined features from our self-constructed Uyghur sentiment corpus at different scales, building upon traditional n-gram features. Our approach includes both traditional bag-of-words features and semantic features that consider contextual information. Experiments demonstrate that among basic features, unigram features achieve the best classification performance for Uyghur text sentiment classification. Combining unigram features with phrase features that consider contextual semantic information can further enhance classification performance. Among the three combined features investigated, the combination of unigram and generalized bi-tagged features achieves the best performance, improving classification accuracy by 1.78% over unigram features alone.

The phrase features investigated in this paper are extracted based on POS collocation rules, consisting of two words with sequential and adjacent relationships. Currently, we cannot perform sentiment classification using units containing more than two words. Future work will focus on investigating how to improve sentiment classification efficiency by expanding phrase feature length and leveraging long-distance dependencies between words. Additionally, we will integrate the features extracted in this paper with word vector features from deep learning models, using them as input features for deep learning models to evaluate their performance in deep learning-based sentiment classification tasks.

References

[1] Liu Bing. Sentiment analysis and opinion mining [C]//Proc of Synthesis Lectures on Human Language Technologies. [S. l.]: Morgan & Claypool, 2012: 152-153.

[2] Pang Bo, Lee L, Vaithyanathan S. Thumbs up?: sentiment classification using machine learning techniques [C]//Proc of Acl-02 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics. 2002: 79-86.

[3] Habernal I, Steinberger J. Supervised sentiment analysis in czech social media [M]. [S. l.]: Pergamon Press, Inc, 2014.

[4] Duwairi Rehab, EI-Orfali M. A study of the effects of preprocessing strategies on sentiment analysis for arabic text [J]. Journal of Information Science, 2014, 40(4): 501-513.

[5] Kim Y. Convolutional neural networks for sentence classification [J]. arXiv: 1408.5882v2, 2014.

[6] Liang Jun, Chai Yumei, Yuan Huibin, et al. Deep learning for Chinese micro-blog sentiment analysis [J]. Journal of Chinese Information Processing, 2014, 28(5): 155-161.

[7] Yierxiati Tuergong, Wushour Silamu, Rexidanmu Tuerhongtai, et al. Construction and analysis of Uyghur emotional corpus [J]. Computer and Modernization, 2017(4): 67-72.

[8] Li Zekui, Zhao Yanyan, Qin Bing, et al. Feature engineering for Chinese microblog sentiment classification [J]. Journal of Shanxi University: Natural Science Edition, 2014, 37(4): 570-579.

[9] Sun Chaohong. Research on micro-blog sentiment classification based on recurrent neural network [D]. Hangzhou: Master' s Thesis of Zhejiang Sci-Tech University, 2017.

[10] Xu Yingying. Research of sentence-level sentiment classification for text based on deep neural network [D]. Shenzhen: Master' s Thesis of Shenzhen University, 2016.

[11] Tian Shengwei, Yu Long, Wang Yuguang. Research on sentiment classification of Uighur reviews [J]. Computer Engineering and Applications, 2011, 47(36): 147-150.

[12] Reyilaimu Paerhati, Meng Xiangtao, Aisikaer Aimudula. Uyghur text sentiment classification based on discriminative keyword model [J]. Computer Engineering, 2014, 40(10): 132-136.

[13] Abudusalamu Dawuti, Yusiyin Yusupu, Aisikaer Aimudula. Emotion recognition from Uyghur sentences based on combinations of class discrimination words and a sentiment dictionary [J]. Journal of Tsinghua University: Natural Science Edition, 2017, 57(2): 197-201.

[14] Li Min, Yu Long, Tian Shengwei, et al. Emotional tendency analysis of Uyghur statement based on deep learning [J]. Computer Engineering and Design, 2016, 37(8): 2213-2217.

[15] Li Dongbai, Tian Shengwei, Yu Long, et al. Deep learning for implicit sentiment classification of Uyghur sentence [J]. Computer Engineering and Design, 2016, 37(9): 2577-2580.

[16] Wang Shuheng, Turgun Ibrahim, Kahaerjiang Aviderexiti, et al. Sentiment classification of Uyghur text based on BLSTM [J]. Computer Engineering and Design, 2017, 38(10): 2879-2886.

[17] Abdukelimu H, Liu Yang, Chen Xinxiong, et al. Learning distributed representations of Uyghur words and morphems [M]. Cham: Springer International Publishing, 2015: 202-211.

[18] Aishan Wumaier, Tuergen Yibulayin. Sentence boundary detection of Uyghur based on rules and statistics [J]. Computer Engineering and Applications, 2010, 46(14): 162-165.

[19] Litip Tohti. The reference grammar of modern Uyghur language [M]. Beijing: China Social Sciences Press, 2012.

[20] Maierhaba Aili, Jiang Wenbin, Wang Zhiyang, et al. Directed graph model of Uyghur morphological analysis [J]. Journal of Software, 2012, 23(12): 94-100.

[21] Maimaiti M, Wumaier A, Abiderexiti K, et al. Bidirectional long short-term memory network with a conditional random field layer for Uyghur part-of-speech tagging [J]. Information, 2017, 8(4): 157.

[22] Rexidanmu Tuerhongtai, Wushour Silamu, Yierxiati Tuergong. Uyghur text sentiment analysis by combining lexical knowledge with machine learning methods [J]. Journal of Chinese Information Processing, 2017, 31(1): 177-183.

[23] Turney P D. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews [C]//Proc of Annual Meeting of the Association for Computational Linguistics. 2002: 417-424.

[24] Gamon M. Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis [C]//Proc of International Conference on Computational Linguistics. 2004: 841-847.

[25] Joshi M, Penstein-Rosé C. Generalizing dependency features for opinion mining [C]//Proc of the ACL-IJCNLP 2009 Conference Short Papers. 2009: 2577-2580.

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