NCYU at IJCNLP-2017 Task 2: Dimensional Sentiment Analysis for Chinese Phrases using Vector Representations

Jui-Feng Yeh*, Jian-Cheng Tsai, Bo-Wei Wu, Tai-You Kuang

Department of Computer Science & Information Engineering,
National Chia-Yi University,
Chiayi city, Taiwan (R.O.C.)
{ralph, s1033027, s1032881, s1050470}@mail.ncyu.edu.tw

Abstract

This paper presents two vector representations proposed by National Chiayi University (NCYU) about phrased-based sentiment detection which was used to compete in dimensional sentiment analysis for Chinese phrases (DSACP) at IJCNLP 2017. The vector-based sentiment phrase-like unit analysis models are proposed in this article. E-HowNet-based clustering is used to obtain the values of valence and arousal for sentiment words first. An out-of-vocabulary function is also defined in this article to measure the dimensional emotion values for unknown words. For predicting the corresponding values of sentiment phrase-like unit, a vector-based approach is proposed here. According to the experimental results, we can find the proposed approach is efficacious.

1 Introduction

As we known, humanity is composed of rationality and sensibility. In the latest decays, the computational thinking based-on logical inference has been mature than that based on perceptual ones. However, the computation technologies cannot be practical perfectly in real life without perceptual sensibility indeed. Consequently, affective computing is one of the most essential trends in artificial intelligence in the near future. Nowadays, logic inference has been developed and practiced in the real-life in latest decays. Actually, no emotional expression makes the intelligent systems and robots to act like machine, but not human. The goal of affective computing is to compute the issues related to, arises from, or deliberately influences emotions. Only rational interpretation without human-like emotion cannot express the full meaning and intension in human-machine interactions.

As increasing of the number of the network users, the on-line social computing has more and more popular. Therefore, the detection of the user in cyber space is more and more desired (Pang and Lee, 2008). Considering of the data styles, the two major ones are image and text in social media such as Facebook, Twitter, LinkedIn, Pinterest, Instagram and Snapchat. Actually, the emotion sates of the users are detected from their post information (Rosenthal et al., 2015). Hao et al. (2015) proposed that the sentiment is analyzed by the complex network. They also used sentiment modes network which built by the sentiment dictionary for sentiment analysis. The method of building sentiment modes network is converted the abstract posts into sentiment fragments by sentiment dictionary and coarse graining to build the sentiment modes. However, to detect the affective information from the image-based data such as picture is still one of the open questions. We plan to classify the emotion from the text data at this time. Due to the symbolic representation of text, the new computing technologies usually applied in treating the problems in natural language first. From the psychological constructionist views, emotions are experienced when affective stated are made meaningful as specific instances of the emotion categories (Lindquis et al., 2016). Lindquis et al. also proposed a think: putting words into feelings and putting feelings into words. From this view point, we can sure that lan-

guage is corrected with emotion. Therefore, the emotional element embedded in natural language is one of important works about affective computing in the near future.

This paper is organized as follows. Section 2 describes the related works; affective computing and the description about the emotion categories and dimensional valence-arousal space are illustrated. Section 3 the proposed method of vector-based approach is presented with derivation of mathematical formulas. In Section 4, we analyze the performance in experimental results of the proposed approach. Finally, Section 5 will draw the conclusion of this paper.

2   Related Works

Since affective computing is a new trend in computer science, considerable quantity researches are invested within the latest twenty years. As we known, the aim of the affective computing is to obtain and provide the functionality of human-like sensibility, the real-life data is important for system building.

Affective computing means the technologies computes that related to, arises from and deliberately influences emotions. The goal of affective computing is to recognize, express, and generate emotions from environment. Instead on rational computing only, the systems are desired to act like human. Since Turing test, computer scientists have much effort on artificial intelligence especially in detecting, percept and express emotion. Therefore, social rules extended to machine in several media type. For example, computer vision technology can be applied to detect the emotional facial expression. Combined with audio and speech, a multimodal agent with emotional interaction functions can be developed and practically applied in real environment. Enabling communication of emotion is an attracting in human-machine interactions. *"The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions."* said by Marvin Minsky. It is trying to make computer with the capability to deal with the emotional issues in human machine interactions.

About recognition of emotion/sentiment, the goal is to classify the type and degree of emotion that is displayed by users. First, we can divide the emotion classification into two kinds: discrete categories and dimensional models. The former means

the emotion classes are discrete and with different constructs. The latter denotes that emotions should be described as continuous value regression in dimensional basis in groups.

Due to the alteration of word usage, Li et al. (2017) pointed out the importance of partition sentence correctly which influence the meaning of the sentiment. The article proposed a method is according the enhanced mutual information (EMI) score of new word user-invented to get the new words. EMI score of the words represent the possibility it is a new words, it the higher the more possible. Similarly, new sentiment word detection is still one of essential issues for affective computing. Herein, Huang et al. (2014) found the sentiment word and polarity prediction of new sentiment word is one of the main contributions of this paper. They used the likelihood ratio test (LRT) method which can quantify the degree of association to find the new words and the method produce the framework is fully unsupervised. Moreover, the polarity prediction of sentiment word is benefit for classification of sentiment Words and sentiment analysis.

Word embedding for sentiment classification is one of main steams (Tang et al., 2014; Yao and Li, 2016). Wang et al. (2016b) adopted the deep learning method, CNN-LSTM, to deal with the sentiment analysis problems. Wang et al. (2016) adopted the deep learning method, CNN-LSTM, to deal with the sentiment analysis problems. Yao and Li (2016) presented the method which finding the feature vector of special word in context is by word2vec. In addition, the paper proposes a new idea which is using the feature vector represent the words' sentiment polarity. Yet, this idea remains more and more research to prove because their corpus in a particular context is limited. Yeh et al. (2016) adopted the linear regression for the values about valence and arousal of Chinese words. Mellem et al. (2016) invested a sentence processing in anterior superior temporal cortex shows a social-emotional bias.

3   Vector Representations and Estimation

The term "sentiment phrase-like unit" is defined as a constitution that is composed of a head word and modifier words. Herein, the head word must be a sentiment word. The modifier can be empty, or degree and negation words. An example is illustrated in the Figure 1. After word seg-

mentation and parsing, a parse tree is obtained. We can find that sentiment word is "喜歡(like)" and two modifiers "不(dis)" and "很(very)" in this sentence. According to the previous definition, three "Sentiment phrase-like unit", "喜歡(like)," 不喜歡(dis-like), "很不喜歡(dis-like very much)" are extracted here.
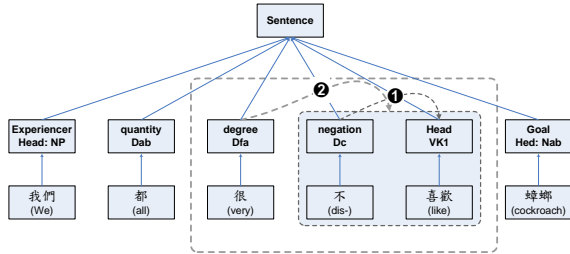

Figure 1: An example for sentiment phrase-like unit in a parse tree of the Chinese sentence "我們很不喜歡蟑螂(We all dis-like cockroach very much)."

Since the sentiment phrase-like unit is essential for affective computing, the question which has been touched from time to time but not explored is how to provide the values of the sentiment phrase-like unit automatically in valence and arousal. Thus, we will describe the process of building a Chinese emotional dictionary with valence-arousal values and the way of how to predict the valence-arousal values of the phrases based on the values of emotional words. In the part of building our dictionary, we used up to four ontologies, E-HowNet, ANTUSD, SentiWordNet 3.0 and CVAW as knowledge bases. In the part of predicting the valence-arousal values of the phrases, we used the CVAP ontology as our training data to train the influence degree value of degree and negative adverb.

The knowledge bases used in the proposed method are illustrated as follows. E-HowNet is a frame-based and extended from HowNet. The purpose of E-HowNet is that makes the concept of the real world can be written in a way that the computer can read. It includes the element, basic concept element, related function and additional symbol. We used these relations of E-HowNet to classify the training words, as shown in Figure 2. The ANTUSD contains 27,370 words and collects handbook noted of word for a long time. It includes NTUSD, NTCIR MOAT task dataset, Chinese Opinion Treebank, AciBiMA, CopeOpi and E-HowNet.SentiWordNet 3.0 is a lexical resource publicly available for research purposes and explicitly devised for supporting sentiment analysis

and opinion mining applications. It is also the result of automatically annotating all WordNet synsets according to their degrees of positivity, negativity, and neutrality. The sentiment scores of SentiWordNet 3.0 are ranging from 0 to 1 including 0 and 1.The CVAW which contains 1,653 words noted with valence-arousal ratings built by Chinese valence-arousal resources. It contains seven attributes which is the number, Valence_Mean, Valence_SD, Arousal_Mean, Arousal_SD and Frequency in sequence. The valence-arousal ratings of CVAW are scored a value between 0 and 9.The CVAP contains 2,251 pharses which classified as six types including degree, degree_negative, mod, mod_degree, negative and negative_degree that annotated with valence-arousal ratings. The valence-arousal ratings of CVAP are scored a value between 0 and 9.

For building a lexical-based sentiment word set, an E-HowNet-based clustering is used here. The word of included relation was defined in the E-HowNet ontology. We cluster the words by using level definition of E-HowNet and let the words of synsets are clustered a group. Every group of synsets are corresponded to a hypernym. At the
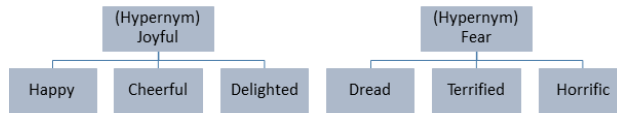

Figure 2: An example of the words in same synset and the same hypernym.

end, the group of hypernym is formed by gathered the words of CVAW gathered, as shown in Figure 2. By using the hypernym, we can get more the emotional words of classified in the same synset. And then, we can expand the words of the dictionary when we used the more hypernym to get more synsets. Based the all above ontology, we convert the valence-arousal values of our dictionary from the valence-arousal values of words. The function of calculate the words is used our dictionary to find the valence-arousal values of words and if the word is build in our dictionary by ANTUSD ontology, we used OOV function to predict the arousal values because the all values of the words are neutral. If the words do not exist in our dictionary, we still can predict the values by the OOV function which is in the next section.

The Sentiment phrase-like unit is composed of the degree adverb, negative adverb and the emo-

tional words. If we want to predict the value, we can start from the value of the words. The valence and arousal value of words can be obtained by the above function and the degree and negative adverb can be acquired from training data, so we used as follow formula to achieve goals:

$$E_v' = D_V * (E_v - C) + C\,, E_v \in Valence \qquad (1)$$

and

$$E_a' = D_A * (E_a - C) + C\,, E_a \in Arousal \qquad (2)$$

where E_v and E_a denote the original value of valence and arousal and E_v' and E_a' represent the value after calculating. C is center value of the valence and arousal and D_V、D_A is a real-number show the degree and negative adverb influence how much the degree of the word value. As the Formula, we believe the modified mood is not the addition result of the composed of words so we used the multiplication to show that. The reason why we need to subtract the C from the E_v、E_a is the subtracted value represent the degree of original emotion and we used that to predict can get the closer emotion value. Additionally, a function for out of vocabulary is defined here. OOV is the abbreviation of the Out of vocabulary and the timing of use is a word cannot be calculated by above function that word does not be found in our dictionary. By OOV function, we can solve the problem which the word does not exist in our dictionary and the formula is as follows.

$$X = \frac{\sum_{Seq_n=1}^{n} \frac{\sum_{i=1}^{N} x_i}{N}}{Seq_n}, X \in Valence \qquad (3)$$

and

$$Y = \frac{\sum_{Seq_n=1}^{n} \frac{\sum_{i=1}^{N} y_i}{N}}{Seq_n}, Y \in Arousal \qquad (4)$$

The OOV function is used to predict valence-arousal value of the undefined words by defined word in our dictionary. We get the average valence-arousal value of each character separately and add the all values to gain the average as the valence-arousal value of the undefined words. Although the OOV function is mainly to solve the above problem, we have also designed a function in it which is support the calculation of phrases. In order to solve the tough which the calculation of phrase may go wrong because the phrase is not consist by only a degree adverb, a negative adverb and a word, we decided adding the function of again searching the adverb and again calculating assume the adverb was found in OOV to let the result of phrase calculation better.

4 Experimental Results

For evaluating for the proposed approach, a system was developed. The data preparation, metric and results are illustrated as following sections. Based on Chinese valence-arousal words 2.0 (CVAW 2.0) developed by (Yu et al. 2016), we have the adjustment set defined in (Yeh et al. 2016). According to the definitions, we proposed two vector-based approaches for predicting for the values of the valence and arousal. Since the test corpus consists 750 word-level and 750 phrases-like level sentiment units. We need the data structure to simultaneously represent the valence-arousal value and the degree of the emotion, the float scores are used here. For a given input test pattern either in word level or phrase-like level, we can get the float scores from 1 to 9 for both valence and arousal which is the mainly dimensional representation about the extent of the emotion from most negative to most positive for valence and from most calm to most excited for arousal. The input and output formats are "word/ phrase_id, word/phrase", and "word/ phrase_id, vallence_value, arousal_value" separately. The following is an example of the input and output, the part of word is :認真(earnest), 寬恕(forgive), and 說服(convince) and the part of phrase is:十分不滿(very dissatisfied), 很可怕(very scary), and 十分壯觀(very spectacular). The word input are "1001, 認真," "1002, 寬恕," and "1003, 說服" and the phrase input are "2001, 十分不滿," "2002, 很可怕," and "2003, 十分壯觀." The word output are "1001, 7.0, 5.0," "1002, 6.2, 4.6," and "1003, 5.2, 3.6." The corresponding phrase output are "2001, 2.256, 7.444," "2002, 2.500, 7.000," and "2003, 7.722, 7.138." There are totally 750 testing instances of test data including word and phrase level units as described previously. The performance is assessed by censoring the difference between machine-predicted ratings and human-annotated which valence and arousal are treated independently. The metrics are the same as those were used in IALP 2016 including Mean Absolute Error (MAE) and Pearson correlation coefficient (PCC). The corresponding formulates are illustrated as the eq. (3) and (4).

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |A_i - P_i| \qquad (5)$$

and

$$r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{A_i - \bar{A}}{\sigma_A}\right)\left(\frac{P_i - \bar{P}}{\sigma_P}\right), \qquad (6)$$

where A_i denotes the human-annotated ratings and P_i denotes the machine-predicted ratings. n is the number of test samples, $\bar{A}$ and $\bar{P}$ respectively denote the arithmetic mean of A and P, and σ is the standard deviation.

According to the above function, we predict the valence-arousal values with the test data and the results are displayed on the dimension graph as Figure 3, Figure 4, Figure 5, and Figure 6. X-axis and Y-axis represents valence and arousal individually. We performed the experiment 2 times by two different proposed approaches to predict the valence-arousal values for the input test patterns. First way is just using OOV function and second way is using above all functions. From the observations, we can get total 1,500 results of valence-arousal values which there are 750 respectively belong to word and phrase and scale to the emotional space.
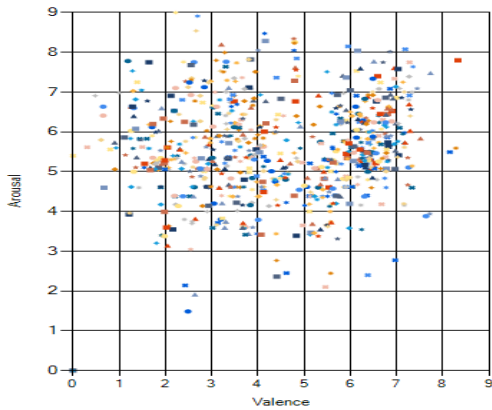

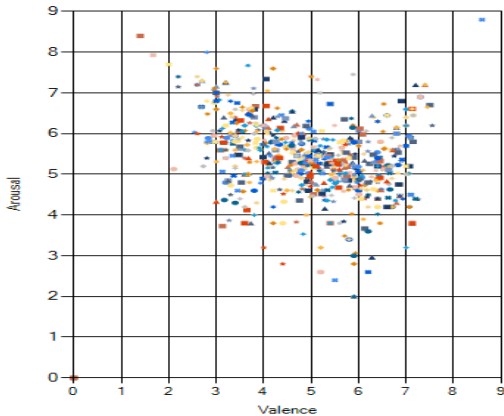Figure 5: The result of NCYU-Run2's phrase
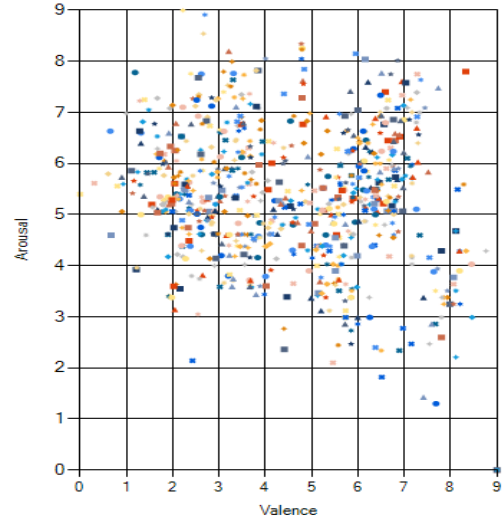

Figure 6: The result of NCYU-Run2's word

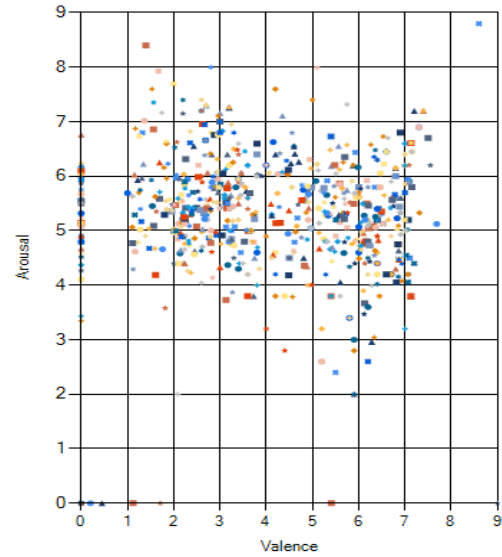Figure 3: The result of NCYU-Run1's phrase

Figure 4: The result of NCYU-Run1's word

Considering of the baseline system, the experimental results about the proposed approaches are shown in following Table 1, Table 2, and Table 3. Table 1 and Table 2 shows the results of analyzed the valence and arousal in MAE and PCC. We can find the performance of the proposed approaches is near to that of baseline system. These results shows that the proposed method is able to provide a correct automatically labeling. From Table 3, we can find that the similar condition appears in mean rank.

116

Submission	Valence MAE (rank)	Valence PCC (rank)
NCYU-Run1	0.9785 (19)	0.685 (19)
Baseline	1.0175 (20)	0.6265 (22)
NCYU-Run2	1.2050 (23)	0.6665 (20)

Table 1: The results of analyzed the valence in averaging performance.

Submission	Arousal MAE (rank)	Arousal PCC (rank)
Baseline	0.819 (14)	0.593 (13)
NCYU-Run1	0.945 (20)	0.549 (16)
NCYU-Run2	0.989 (22)	0.534 (21)

Table 2: The result of analyzed the arousal in averaging performance

Submission	Mean Rank
Baseline	17.25
NCYU-Run1	18.5
NCYU-Run2	21.5

Table 3: The mean rank of averaging performance

5 Conclusions

A vector-based approach containing the E-HowNet-based clustering and OOV function for sentiment phrase-like unit is proposed in this paper. Since the dimensional affective representation is increasing, a valence-arousal space is reasonable. According the structure of sentiment phrase-like unit that contains the emotion word as head-word and negation and degree words as modifiers is defined. Applied the proposed method, we can observe that the proposed approach is workable and efficient.

References

Hao, X., An, H., Zhang, L., Li, H., and Wei, G. 2015. Sentiment Diffusion of Public Opinions about Hot Events: Based on Complex Network. *PloS one*, 10(10), e0140027.

Huang, M., Ye, B., Wang, Y., Chen, H., Cheng, J., and Zhu, X. 2014. New Word Detection for Sentiment Analysis. In *ACL (1)* : 531-541.

Li, W., Guo, K., Shi, Y., Zhu, L., & Zheng, Y. 2017. Improved New Word Detection Method Used in Tourism Field. *Procedia Computer Science*, 108: 1251-1260.

Lindquist, K., Gendron, M., Satpute, A. B., & Lindquist, K. 2016. Language and emotion. In *Handbook of emotions*. The Guilford Press New York, NY.

Mellem, M. S., Jasmin, K. M., Peng, C., & Martin, A. 2016. Sentence processing in anterior superior temporal cortex shows a social-emotional bias. Neuropsychologia, 89: 217-224.

Pang, B. and Lee L. 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1–135.

Rosenthal S., Nakov P., Kiritchenko S., Mohammad Saif M, 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: 451–463.

Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T., & Qin, B. 2014. Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification. In *ACL (1)* : 1555-1565.

Wang, J., Yu, L. C., Lai, K. R., & Zhang, X. 2016. August). Dimensional sentiment analysis using a regional CNN-LSTM model. In ACL 2016— *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*. Berlin, Germany 2: 225-230.

Yao, Y., & Li, G. 2016. Context-aware Sentiment Word Identification: sentiword2vec. arXiv preprint arXiv:1612.03769.

Yeh, J. F., Kuang, T. Y., Huang, Y. J., & Wu, M. R. 2016. Dimensional sentiment analysis in valence-arousal for Chinese words by linear regression. In *Proceedings of International Conference on Asian Language Processing (IALP)*, 2016:328-331

Yu L. C., Lee L.-H., Hao S., Wang J., He Y., Hu J., Lai K. Robert, and Zhang X.. 2016. Building Chinese affective resources in valence-arousal dimensions. In *Proceedings of NAACL/HLT-16*: 540-545.