4.3 Results of Using CVAW to Predict the VA Ratings of CVAT

To demonstrate the application of the constructed affective resources, this experiment adopted a simple aggregate-and-average method (Taboada et al. 2011) to predict the VA ratings of the CVAT corpus using CVAW words. In this approach, the valence (or arousal) rating of a given sentence was calculated by averaging the valence (or arousal) ratings of the words matched in the CVAW in that sentence. Once the predicted values of the VA ratings for the sentences were obtained, they were compared to the corresponding actual values in the CVAT to calculate MAE, RMSE and Pearson correlation coefficient r, as shown in Table 2. Notice that the sentences which contain no affective words in the CVAW were not included for performance calculation (herein 30 sentences). The results using ANEW to predict the VA rating of 20 English forum discussions were also included for comparison (Paltoglou et al., 2013).

The results show that the average tokens of the CVAT sentences are around 35 which is much smaller than those of the English forum discussions (long texts). Both English and Chinese resources had a similar error rates (MAE and RMSE) for valence, while the English resource outperformed the Chinese resource in terms of arousal rates. In addition, both the English and Chinese resources had a lower correlation for arousal than for valence, indicating again that the arousal dimension is more difficult to predict. Table 2 also shows the performance for each category in CVAT. For valence, Laptop achieved the lowest error rate, while News and Hotel had a higher correlation. The respective ranges of MAE, RMSE and r are 0.95~1.48, 1.21~1.77 and 0.30~0.61. For arousal, Book yielded the lowest error rate, while Hotel and Book yielded a better correlation. The respective ranges of MAE, RMSE and r are 0.89~1.11, 1.11~1.40 and 0.04~0.22.

5 Conclusions and Future Work

This study presents a Chinese affective lexicon with 1,653 words and a corpus of 2,009 sentences with six different categories, both annotated with valence-arousal values. A corpus cleanup procedure was used to remove outlier ratings and improper texts to improve quality. Experimental re-sults provided a feasibility evaluation and baseline performance for VA prediction using the constructed resources. Future work will focus on building useful dimensional sentiment applications based on the constructed resources.

References

Carmen Banea, Rada Mihalcea, and Janyce Wiebe. 2013. Porting multilingual subjectivity resources across languages. *IEEE Trans. Affective Computing*, 4(2):211-225.

Margaret M. Bradley and Peter J. Lang. 1999. Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical Report C-1, University of Florida, Gainesville, FL.

Margaret M. Bradley and Peter J. Lang. 2007. Affective Norms for English Text (ANET): Affective ratings of text and instruction manual. Technical Report D-1, University of Florida, Gainesville, FL.

R. A. Calvo and Sidney. D'Mello. 2010. Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Trans. Affective Computing*, 1(1): 18-37.

Paul Ekman. 1992. An argument for basic emotions. *Cognition and Emotion*, 6:169-200.

Andrea Esuli and Fabrizio Sebastiani. 2006. Senti-WordNet: a publicly available lexical resource for opinion mining. In *Proc. of LREC-06*, pages 417-422.

Ronen Feldman. 2013. Techniques and applications for sentiment a0nalysis. *Communications of the ACM*, 56(4):82-89.

Didem Gökçay, Erdinç İşbilir and Gülsen Yıldırım. 2012. Predicting the sentiment in sentences based on words: an exploratory study on ANEW and ANET. In *Proc. of CogInfoCom-12*, pages 715-718.

C.-L. Huang, C. K. Chung, N. Hui, Y.-C. Lin, Y.-T. Seih, W.-C. Chen, B. Lam, M. Bond, and James W. Pennebaker. 2012. The development of the Chinese Linguistic Inquiry and Word Count dictionary. *Chinese Journal of Psychology*, 54(2):185-201.

Lun-Wei Ku and Hsin-Hsi Chen. 2007. Mining opinions from the web: beyond relevance retrieval. *Journal of the American Society for Information Science and Technology*, 58(12), 1838-1850.

Peter J. Lang. 1980. Behavioral treatment and bio-behavioral assessment: Computer applications. *Tech-*