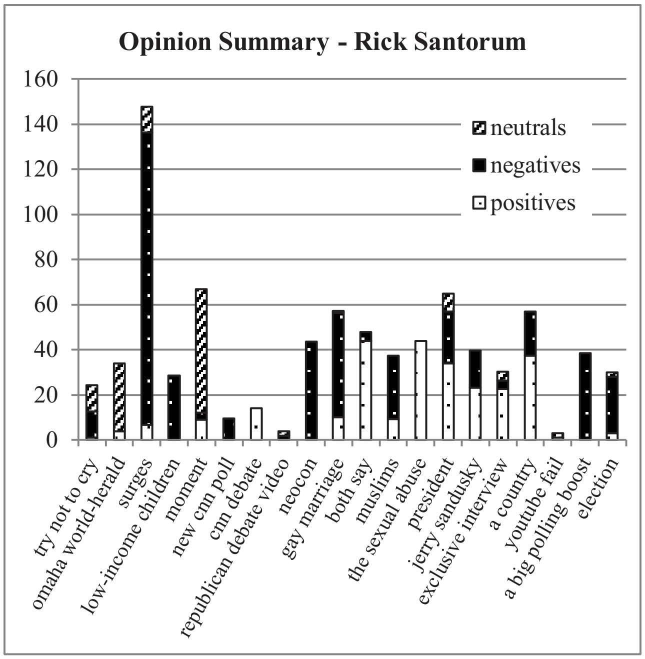

Figure 2: Opinion summary for Rick Santorum

Evaluation and Results

In order to evaluate the performance of the adjusted PMI measure (the product of the average PMI measure and the adjusted count) as constraint on aspect extraction we annotated all extracted noun phrases of two political candidates either as "aspect" or "no aspect".

A noun phrase is labeled as aspect if it represents either a generic political topic, e.g. "foreign policy", or a concrete topic that was relevant for this election's context, e.g. "occupy movement". Classification of noun phrases is based on their constraint score. A higher score means that the noun phrase is more likely to be an aspect. Figure 3 compares the performance of the adjusted PMI measure to pure frequency score as constraint on aspect extraction.

The lift chart in Figure 3 visualizes classification performance depending on the number of included noun phrases in Figures 1 and 2. Note that in the critical region located between 0 and 3 percent noun phrase ratio, where the highest scoring noun phrases are located, adjusted PMI measure correctly classifies more aspects than frequency-based scoring. Later, we set the threshold of included noun phrases to 20. Table 3 justifies this setting, as it presents average classification accuracies that were calculated on two different data sets (Rick Perry and Mitt Romney) with varying threshold of included noun phrases. In our case, the adjusted PMI measure as constraint achieves highest average accuracy at a threshold of 20. These results can be interpreted as follows: The PMI adjustment weights out the score of some of the frequent phrases that are, although high frequency-based score, no aspects, and tries to give low-frequency aspects a scoring boost. This leads to more accurate extraction results than pure frequency scoring. Additionally, this implies that the meronymy relationship between politicians and their campaign topics holds.

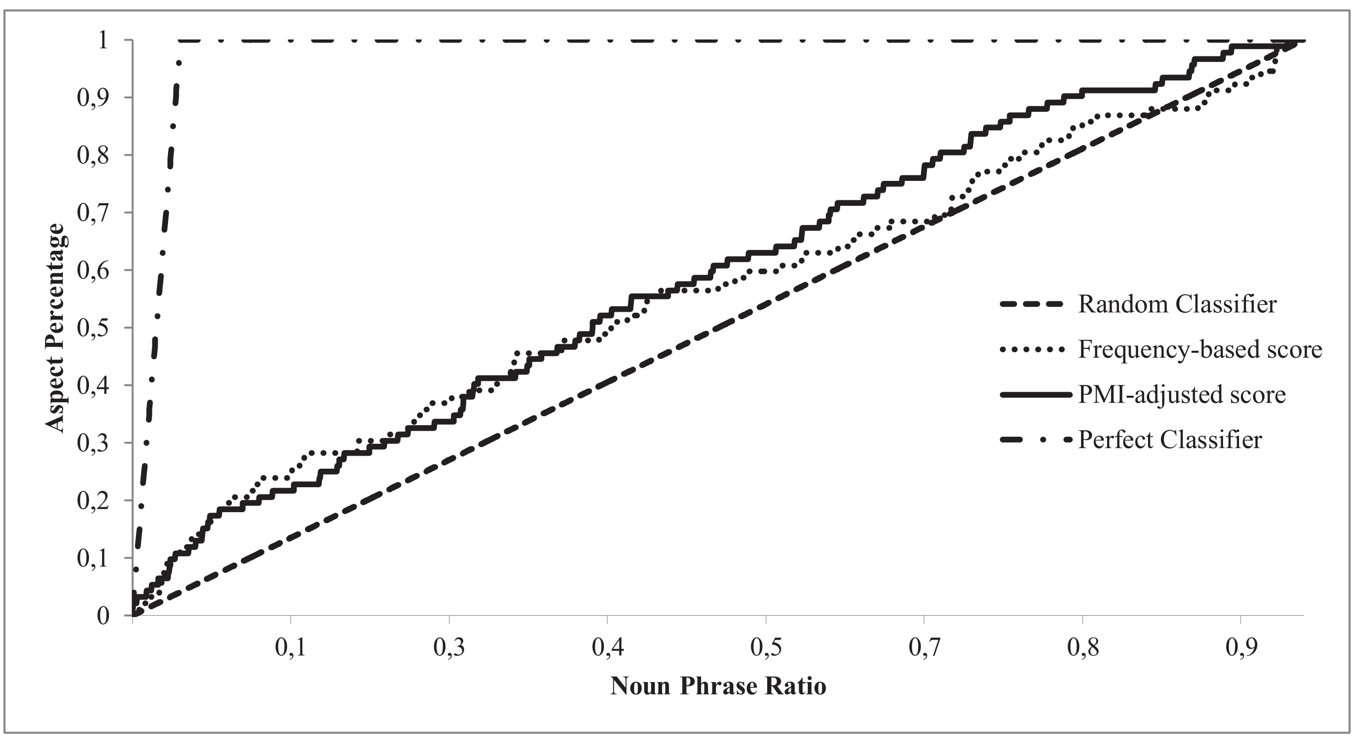

Figure 3: Lift chart for classification performance

Threshold	Average accuracy	
	Adjusted PMI measure	**Frequency score**
20	**22,5%**	17,5%
40	20,25%	16,25%
60	17,50%	18,34%
80	16,75%	**20,00%**
100	16,00%	19,50%

Table 3: Average accuracies

Conclusions and Future Work

This paper presents the challenging task of aspect-based opinion summarization on Twitter data in the domain of politics, which falls into the application category of social media monitoring.

It was discussed that although Twitter data can easily be gathered, special considerations in retrieval and pre-processing are needed. NLTK's built-in pre-processing functionalities were found to be not completely sufficient for informal text corpora. We extracted relevant aspects with a newly introduced combination of the PMI measure and phrase frequency as constraint. Aspect extraction and pruning methods presented in this paper can be applied in any domain where a meronymy relationship of opinion targets and aspects holds true. We verified this relationship for political candidates and their campaign topics. The evaluation of the PMI adjusted measure as constraint on aspect extraction shows that the meronymy relationship between politicians and their campaign holds.

Possibilities for future work include the learning of other domain-specific opinion words like nouns and verbs. Such a classification task would probably need to involve syntactic dependencies as features. Both time and regional distinctions could reveal trends and allow a more detailed presentation of political topics and associated sentiment. This could reveal that a certain topic causes positive reactions in one state, while it gets mostly negative comments in another state. In terms of aspect-level sentiment, the simple distance-weighted score presented here can be improved when it is assured that particular opinion words are expressed in relation to the aspect or the opinion target. Sophisticated analysis of long-distance opinion shifter dependencies are expected to increase the reliability of aggregated aspect sentiment.

References

Brody, S. & Diakopoulos, N., 2011. Using Word Lengthening to Detect Sentiment in Microblogs. *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pp. 562-570.

Davidov, D., Tsur, O. & Rappoport, A., 2010. Enhanced Sentiment Learning Using Twitter Hashtags and Smileys. *Proceeding of the 23rd international conference on Computational Linguistics (COLING)*.

Diakopoulos, N. & Shamma, D., 2010. Characterizing Debate Performance via Aggregated Twitter Sentiment. *Proceedings of the 28th int. conference on Human factors in computing systems*.

Go, A., Bhayani, R. & Huang, L., 2009. Twitter Sentiment Classification using Distant Supervision.

Hatzivassiloglou, V. & McKeown, K., 1997. Predicting the Semantic Orientation of Adjectives. *Proceedings of the Joint ACL/EACL Conference*, pp. 174-181.

Hu, M. & Liu, B., 2004. Mining and Summarizing Customer Reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*.

Hu, M. & Liu, B., 2006. Opinion Feature Extraction Using Class Sequential Rules. *Proceedings of the Spring Symposia on Computational Approaches to Analyzing Weblogs*.

Kim, S.-M. & Hovy, E., 2004. Determining the Sentiment of Opinions. *Proceedings of the International Conference on Computational Linguistics*.

Liu, B., 2010. Sentiment Analysis and Subjectivity. In: *Handbook of Natural Language Processing.* Boca Raton, FL: CRC Press, pp. 627 - 666.

Manning, C. & Schütze, H., 1999. *Foundiations of Statistical Natural Language Processing.* Massachusetts Institue of Technology - Cambridge: THE MIT PRESS.

Pak, A. & Paroubek, P., 2010. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. *Proceedings of the Seventh conference on International Language Resources and Evaluation*.

Pang, B. & Lee, L., 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*.

Pang, B., Lee, L. & Vaithyanathan, S., 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

Popescu, A.-M. & Etzioni, O., 2005. Extracting Product Features and Opinions from Reviews. *Proceedings of Conference on Empirical Methods in Natural Language*.

Qiu, G., Liu, B., Bu, J. & Chen, C., 2011. Opinion Word Expansion and Target Extraction using Double Propagation. *Computational Linguistics Vol. 37*.

Riloff, E. & Wiebe, J., 2003. Learning Extraction Patterns for Subjective Expressions. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

Ringsquandl, M. & Petkovic, D., 2012. Expanding General Lexicon with Domain Specific Opinion Words Using Semi-Supervised Approach. *BRACIS 2012 - WTI, IV International Workshop on Web and Text Intelligence*.

Turney, P., 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *Proceedings of the Association for Computational Linguistics*, p. 417–424.

Wiebe, J., Wilson, T. & Cardie, C., 2005. Annotating Expressions of Opinions and Emotions in Language. *Language Resources and Evaluation*, pp. 164-210.

Wilson, T., Wiebe, J. & Hoffmann, P., 2005. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing*.