Analyzing Political Sentiment on Twitter

Martin Ringsquandl Dušan Petković

University of Applied Sciences, Rosenheim, Germany
martin.ringsquandl@googlemail.com
petkovic@fh-rosenheim.de

Abstract

- Microblogs pose a great challenge to aspect-based opinon summarization, including sentiment classification and aspect extraction
- Political campaigns can leverage from opinions in online informal political discourse, especially in aggregated form
- We propose a new method for aspect extraction based on Pointwise Mutual Information (PMI) and adjusted term frequency
- Evaluation shows that the meronymy relationship between politicians and their topics holds and improves accuracy of aspect extraction

1.Introduction

- Searching for people's opinions via surveys and polls has been an expensive and time-consuming task
- Due to the vast amount of user-generated content, there is a growing need for computational processing of sentiment analysis
- Systems dealing with Microblogging services like Twitter and other social communication platforms are still in the early stage of development, although many applications, including politics and marketing campaigns, can leverage from the use of social media and especially Twitter
- Detailed opinion summaries are necessary, such as distributions of opinions by topic or entity
- We propoose the application of the the Pointwise Mutual Information (PMI) measure for aspect extraction in the domain of online informal political discourse
- We generate aspect-based opinion summaries for each politician
- We show that the meronymy relationship between politicians and their campaign topics holds and adds valuable information to aspect extraction

2. Data Colletcion and Preprocessing

- Candidates of the Republican presidential primaries in the USA
- Twitter was queried several times in November and December of 2011 for tweets mentioning the names of the candidates

Table 1. Republican presidential candidates and the number of corresponding tweets

Candidate	Number of tweets
Herman Cain	7204
Jon Huntsman	5741
Mitt Romney	7243
Michele Bachmann	6864
Newt Gingrich	7199
Rick Santorum	5932
Ron Paul	7279
Rick Perry	7266

- Natural Language Processing was done in Python and NLTK
- Tokenizaion: Punkt Sentence Tokenizer and Treebank Word Tokenizer
- Penn Treebank POS tagging and regular expression noun phrase chunking
- Explicit aspects are assumed to be noun phrases (Implicit aspects are ignored)
- Noise Cleaning:
  i. Hyperlinks and hashes (tags were preserved)
  ii. Usernames with preceding "@" which expresses a reply to another user
  iii. The "RT" keyword which is used to indicate that the following tweet is a retweet

3.Aspect Extraction

- The goal is to find relevant campaign topics (aspects) which are assumed to form a meronymy with their associated political candidates
- The PMI measure and phrase frequency form a constraint for aspect extraction
- According to [3], during aspect extraction (also called feature extraction) it is assumed that only noun phrases are relevant aspects of the opinion targets
- Manning and Schütze in [17] argue that sparseness is a difficult problem for PMI, because bigrams composed of low-frequency words will receive a higher score than those composed of high-frequency words. (The notion of Pointwise Mutual Information has been introduced by Dunning [18].) This problem is compensated in our work, as can be seen in the next subsection
- Equation 1 shows the simplified version of the PMI measure used to search the web for the number of hits of phrases x and y, which is often called PMI-IR (Information Retrieval):

$$PMI_{IR}(x,y) = \frac{hits(x\ NEAR\ y)}{hits(x)hits(y)}$$

(Eq. 1)

Example: PMI measure for candidate Mitt Romney and the aspect "teaparty" is calculated for each discriminator d in the set {of, 's, about} and is given in Equation 2:

$$PMI_{IR}(Mitt\ Romney + d, teaparty) = \frac{hits(Mitt\ Romney + d\ NEAR\ teaparty)}{hits(Mitt\ Romney + d)hits(teaparty)}$$

(Eq. 2)

3.1 Constraint on Aspect Extraction

- PMI spareseness problem: bigrams composed of low-frequency words score higher than high-frequency bigrams
- Can be compensated in conjunction with noun phrases' frequency
- Retweets are of less value and assumed to spread exponentially (Equation 3)

$$c_{adjust} = (c - c_{rt}) + \ln(c_{rt})$$

(Eq. 3)

- Top resulting aspects of each candidate are meassured by the product of the average PMI measure (PMI_{avg}) and the adjusted count of tweets c_{adjust}

Table 2: Mitt Romney: Aspect extraction results

c	c_{rt}	*Aspect*	c_{adjust}	$PMI_{avg} \times c_{adjust}$	c_{pos}	c_{neg}	c_{neu}
36	19	planned parenthood	19.94	4.69029	6.00	13.94	0
218	96	teaparty	126.56	4.00904	48.76	47.14	36.40
36	20	boston globe	19.00	3.76347	15.71	4.61	0
16	3	the character of his opponents	14.10	3.19344	0	14.10	0
77	7	campaign ad	71.95	3.16766	1.00	69.95	1.00
11	4	an equality advocate out	8.39	2.86077	0	8.39	0
153	80	mitt2012	77.38	2.40483	29.40	27.50	25.83
67	6	Defends	62.79	2.14401	2.00	60.79	0
9	2	job prospects	7.69	2.04178	7.69	0	0
55	5	obama quote	51.61	2.00457	0	51.61	0
9	8	a repeat	3.08	1.97861	0	3.08	0
52	22	sesame street	33.09	1.82636	18.39	15.89	0
96	43	a beer	56.76	1.82251	0	56.76	0
25	0	Reuters	25.00	1.81862	11.00	14.00	0
68	23	a cigarette	48.14	1.75517	0	48.14	0
56	6	president Obama	51.79	1.68823	11.10	41.10	0
59	22	Wants	40.09	1.67530	9.71	30.95	1.00
9	1	cnn poll	8.00	1.66771	4.00	0.00	4.00
76	30	company showed profits	49.40	1.64969	0	49.40	0
10	0	washington post	10.00	1.54782	0	9.00	1.00

- Context should be provided for every aspect
- For example:
  "IT'S ON: Democrats Are Waging War Over Romney's 'Sleazy' Campaign Ad http://..."
- Most of the tweets are messages containing headlines of online news articles posted by different users.
- This leads to biased counts and unreliable results

4. Sentiment Classification

- Two steps in unsupervised sentiment classification:
  i. Assemble a general or domain dependent opinion lexicon for words or opinion phrases
  ii. Classify sentiment based on a statistical measure

4.1Building Lexicon

- Subjectivity clues lexicon, presented in [19], was used to detect semantic orientation at word-level
- Consists of 2296 positive, 4138 negative and 444 neutral distinct opinion words
- Highly informal domains like tweets or social networks communication need domain specific lexicons
- In [20], we extract adjectives from the Twitter corpus and expand the general lexicon based on the idea of sentiment consistency of Hatzivassiloglou and McKeown [21]

4.2 Word-Level Sentment

- Semantic orientation of word w is the class which maximizes the probability c conditional on w, where C= {positive, negative, neutral} and $c \in$ C (Equation 4)
- Every word w can be represented as the set of its synonym retrieved from WordNet

$$SO(w) = \arg\max_{c \in C} P(c|w) = \arg\max_{c \in C} \frac{\sum_{i=1}^{n} count(syn_i, c)}{|synset_w|} P(c)$$

(Eq. 4)

- Directly preceding negation changes semantic orientation from negative to positive and vice versa
- Neutral sentiment is not affected by negations

4.3Aspect-Level Sentiment

- Sum up the semantic orientation of all words in sentence s that mentions aspect a, weighted relative to its distance to the aspect
- Equation 5 shows how the aspect-level sentiment score is calculated:

$$score(a,s) = \sum_{w_i \in s} \frac{SO(w_i)}{dist(w_i,a)}$$

(Eq. 5)

- **score(a,s)** > 0 means that a sentiment about the aspect is positive, **score(a,s)** < 0 means negative, **score(a,s)** = 0 is neutral

5. Opinion Summaries

- Aggregated aspect-level sentiment for each candidate is visualized in form of a bar chart
- Aspects like "teaparty" and "mitt2012" which are often used as hashtags, exhibit uniform class distribution of all polarity classes

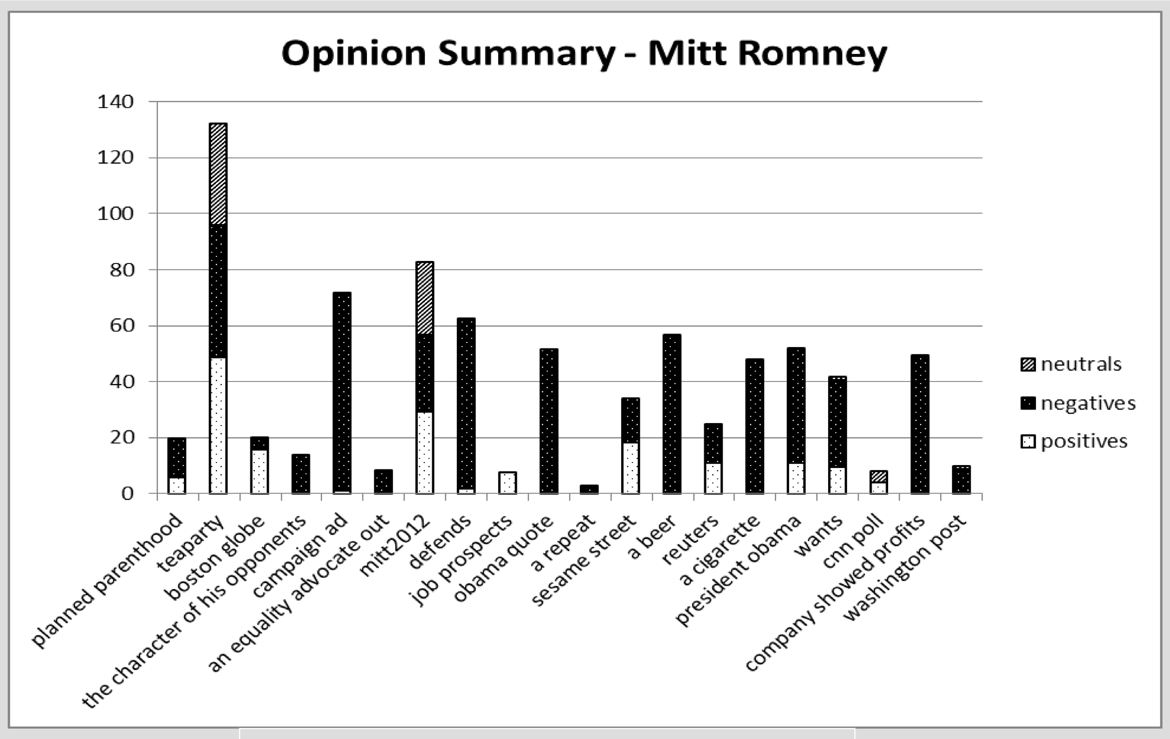

Figure 1: Opinion summary for Mitt Romney

6. Evaluation and Results

- Performance of the adjusted PMI measure as constraint on aspect extraction
- A noun phrase is labeled as aspect 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
- 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, depending on the number of included noun phrases (critical region located between 0 and 3 percent noun phrase ratio)
- Table 3 presents average classification accuracies that were calculated on two different data sets (Rick Perry and Mitt Romney) with varying threshold of included noun phrases
- **Interpretation**: 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
  i. more accurate extraction results than pure frequency scoring
  ii. implies a meronymy between politicians and their campaign topics

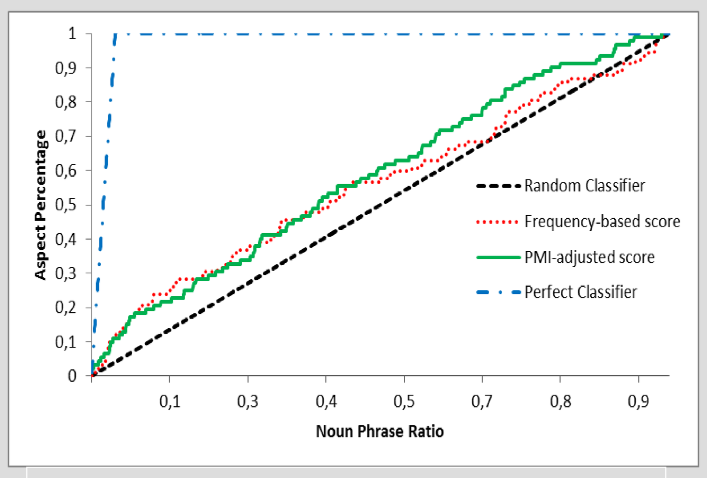

Figure 3: Lift chart for classification performance

Threshold	Average accuracy Adjusted PMI	Frequency score
20	**22.50%**	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

7. Conclusions and Future Work

- Paper presents the challenging task of aspect-based opinion summarization on Twitter data (social media monitoring) in the domain of politics
- Twitter data can easily be gathered, but special considerations in retrieval and pre-processing are needed
- NLTK's built-in pre-processing functionalities not completely sufficient for informal text corpora
- Newly introduced combination of the PMI measure and phrase frequency as constraint on aspect extraction (can be applied in any domain where a meronymy relationship of opinion targets and aspects holds true)
- Evaluation shows that the meronymy between politicians and their campaign topics holds
- Future work:
  i. learning of other domain-specific opinion words like nouns and verbs
  ii. Both time and regional distinctions for more detailed summaries
  iii. Assure that particular opinion words are expressed in relation to the aspect or the opinion target
  iv. long-distance opinion shifter dependencies for aggregated aspect sentiment

References

[1] Turney, P. 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews, in Proc. of the Association for Comp. Linguistics.
[2] Pang; B.; Lee, L.; Vaithyanathan, S. 2002. Thumbs up? Sentiment Classification using Machine Learning Techniques", in Proc. of EMNLP.
[3] Liu, B. 2010. Sentiment Analysis and Subjectivity, in Indurkhya, N.; Damerau, F. - Handbook of Natural Language Processing.
[4] Riloff, E.; Wiebe, J. 2003. Learning Extraction Patterns for Subjective Expressions. Proc. of the Conf. on Empirical Methods in Natural Language Processing.
[5] Wiebe, J.; Wilson, T.; Cardie, C. 2005. Annotating Expressions of Opinions and Emotions, in Language Resources and Evaluation.
[6] Hu, M.; Liu, B. 2004. Mining and Summarizing Customer Reviews. Proc. of the 10th ACM SIGKDD international conference on Knowledge discovery and data mining.
[7] Hu, M.; Liu, B. 2006. Opinion Feature Extraction Using Class Sequential Rules, in Proc. of the Spring Symposia on Computational Approaches to Analyzing Weblogs.
[8] Qiu, G.; Liu, B.; Bu, J.; Chen, C. 2011. Opinion Word Expansion and Target Extraction using Double Propagation, in Computational Linguistics Vol. 37.
[9] Pak, A.; Paroubek, P. 2010 Twitter as a Corpus for Sentiment Analysis and Opinion Mining,in Proc. of the 7th Conf. on Int. Language Resources and Evaluation.
[10] Go, A. 2009. Twitter Sentiment Classification using Distant Supervision, in In CS224N Project Report, Stanford.
[11] Brody, S.; Diakopoulos, N. 2011. Using Word Lengthening to Detect Sentiment in Microblogs, in Proc. of the Conf. on Empirical Methods in Natural Language Processing.
[12] Davidov, D.; Tsur, O.; Rappoport. A. -2010. Enhanced Sentiment Learning Using Twitter Hashtags and Smileys, COLLING.
[13] Mullen, T.; Malouf, R. 2006. A preliminary investigation into sentiment analysis of informal political discourse, in Proc. of AAAI Spring Symposium on Comp. Approaches to Analyzing Weblogs.
[14] Diakopoulos, N.; Shamma, D. 2010. Characterizing Debate Performance via Aggregated Twitter Sentiment, in Proc. of the 28th Int. Conf. on Human factors in computing systems.
[15] Kim, S.; Hovy, E. 2004. Determining the Sentiment of Opinions, in Proc. of the Int. Conf. on Computational Linguistics.
[16] Popescu, A.; Etzioni, O. 2005. Extracting product features and opinions from reviews, in Proc. of HLT/EMNLP.
[17] Manning, C.; Schütze, H. 1999. Foundations of Statistical Natural Language Processing, MIT Press.
[18] Dunning, T. 1993. Accurate methods for the statistics of surprise and coincidence, Computational Linguistics, 19(1).
[19] Wilson, T., Wiebe, J.; Hoffmann, P. 2005. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis, in Proc. of the Human Language Technology Conf. and the Conf. on Empirical Methods in Natural Language Processing.
[20] 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 (to be published).
[21] Hatzivassiloglou, V.; McKeown, K. 1997. Predicting the Semantic Orientation of Adjectives, in Proc. of the Joint ACL/EACL Conf..