Aspect Based Sentiment Analysis for Large Documents with Applications to US Presidential Elections 2016

Completed Research Paper

Raza Ahmad
Lahore University of Management Sciences, Lahore Pakistan
syed.raza@lums.edu.pk

Ahsan Pervaiz
Lahore University of Management Sciences, Lahore Pakistan
ahsan.pervaiz@lums.edu.pk

Humdah Mannan
Lahore University of Management Sciences, Lahore Pakistan
humdah.mannan@lums.edu.pk

Fareed Zaffar
Lahore University of Management Sciences, Lahore Pakistan
fareed.zaffar@lums.edu.pk

Abstract

Aspect based sentiment analysis (ABSA) deals with the fine grained analysis of text to extract entities and aspects and analyze sentiments expressed towards them. Previous work in this area has mostly focused on data of short reviews for products, restaurants and services. We explore ABSA for human entities in the context of large documents like news articles. We create the first-of-its-kind corpus containing multiple entities and aspects from US news articles consisting of approximately 1000 annotated sentences in 300 articles. We develop a novel algorithm to mine entity-aspect pairs from large documents and perform sentiment analysis on them. We demonstrate the application of our algorithm to social and political factors by analyzing the campaign for US presidential elections of 2016. We analyze the frequency and intensity of newspaper coverage in a cross-sectional data from various newspapers and find interesting evidence of catering to a partisan audience and consumer preferences by focusing on selective aspects of presidential candidates in different demographics.

Keywords

sentiment analysis, aspect based sentiment analysis, US presidential elections, entity aspect extraction.

Introduction

Sentiment analysis is a type of subjectivity analysis in which a piece of text is generally labeled as positive, negative or neutral on the basis of its linguistic structure and certain subjective cues (Pang and Lee, 2008). The piece of text could either be a whole document, a paragraph, a sentence or a part thereof. While small sentences could just be bipolar in sentiment, documents having complex structures contain a variety of sentiments, often conflicting, aimed at different entities and topics. Labeling them as merely positive or negative is not suitable for many practical purposes. Hence to conduct a useful and accurate analysis, different aspects of the mentioned entities are extracted from a document to perform sentiment analysis on them.

Aspect based sentiment analysis (ABSA) has received significant attention in the last few years. Much of the research work done in this area has been limited to identifying aspects for a single entity of interest in reviews data. SemEval 2015 and 2016 extend this to reviews containing multiple entities and aspects. However, these reviews are short and are inherently targeted towards explaining different aspects of the reviewed object. Entity extraction in all the previous work has mostly focused on products, movies and other items. Extraction of entities and aspects for human beings remains relatively unexplored.

Most of the textual data in the world is in the form of large text documents that have complex structural, semantic and logical constructs, e.g., books, news articles, research papers, blogs, etc. These documents usually span multiple paragraphs and dozens of sentences with references to many entities along with the discussion of their aspects. Hence mining entities and aspects at this scale and performing sentiment analysis on them still remains an open and challenging problem.

In this paper, we present a novel algorithm to mine multiple entity aspect pairs in large documents for human entities. Our algorithm is a joint model of coreference resolution, semantic role labeling and sentiment classification. We construct a corpus of news articles from different US news outlets which contains around 1000 annotated sentences from approximately 300 articles. Each sentence has been annotated for all the entity-aspect pairs and their corresponding sentiment scores.

The rest of the paper is as follows. Section 2 discusses the related work, section 3 describes the dataset, section 4 presents our algorithm, section 5 details our evaluation criteria and section 6 discusses the results. Finally we analyze the application of our work to US elections 2016 in section 7 and end with conclusion and future directions in section 8.

Related Work

Aspect based sentiment analysis has been an active area of research for the past few years. It consists of two mains tasks: extracting entity-aspect pairs and identifying sentiment for them. To extract entities and aspects, researchers have applied frequency based methods using noun phrases, methods exploiting relations between opinions and targets, supervised and unsupervised learning techniques and topic modeling based approaches (Long et al., 2010; Jiang et al., 2010; Sauper et al., 2011; Liu et al., 2015). Along similar lines, (Kim and Hovy, 2006) extract opinion holder and topic of discussion from isolated sentences in online news media text. Similarly, (Prasojo et al., 2015) work on news comments in the context of ABSA. These works do not extend to long document structures.

Exercises in the International Workshop on Semantic Evaluation (SemEval) have motivated a lot of researchers in this area during the past couple of years. These exercises concern ABSA in datasets of short reviews of different products. SemEval 2014 introduced ABSA for the first time in its Task 4 (Pontiki et al., 2014) for extracting all aspects of a given entity from reviews of restaurants and laptops. SemEval 2015 through 2017 extended upon its predecessor to extract multiple entities and aspects from reviews (Pontiki et al., 2015; Nakov et al., 2016; Rosenthal et al., 2017). The tasks attracted submissions from a number of teams who employed different techniques including topic modeling, sentiment lexicons, MaxEnt classifiers, parse trees, word clustering and deep learning.

More recently, new methods have been developed to improve different tasks in ABSA involving the application of novel neural network architectures. (Wang et al., 2016) propose a modification on long short-term memory (LSTM) for taking aspects into account when determining sentiment. (Poria et al., 2016) developed a very deep convolutional neural network for extracting aspect words from sentences. Similarly, (Liu et al., 2016) improve the aspect extraction phase using semantic similarities and associations between different aspects. (Wang et al., 2016) further the idea of aspect extraction and develop a deep architecture for simultaneous extraction of aspect and opinion terms in a review.

In the light of our analysis, we conclude that almost all the previous work done in this area has circumscribed to short reviews and does not take up the task of dissecting long documents for aspect based sentiment analysis. Moreover, it does not deal with human entities and hence differs in the semantic relation of entities with their aspects. The predominant entities in these reviews are different items like cameras, laptops and restaurants, etc. Hence, we aim to undertake this task of aspect based sentiment analysis for human entities and their aspects in long complex text structures like news articles.

Dataset

Most of the datasets publicly available consist of reviews that are either annotated for multiple aspects of a single entity, do not qualify as large documents or contain annotations for products and services only (McAuley et al., 2012; Pontiki et al., 2015; Nakov et al., 2016). Hence, we construct our own dataset of documents spanning multiple paragraphs which contain human entities with mentions of their aspects. We chose 7 newspapers including *Chicago Tribune, Dallas News, Houston Chronicle, LA Times, NY Post, Washington Post and Seattle Times*. From the top 25 newspapers by circulation count (Peters and Woolley, 2012), we chose these 7 as they covered approximately all major US demographics and contained the most articles relevant to elections. We collected all their articles over a period of one year till 21st September 2015 and excluded the ones not relevant to US elections through topic modeling (Blei et al., 2003) and manual analysis. Our data set[1] finally consisted of around 1000 sentences from 300 articles. We use predefined inventories of entities and aspects with the list of entities comprising 14 presidential candidates from both major political parties. We consider campaign issues as aspects as media mostly discusses the candidates with respect to these campaign issues like foreign policy, economy, etc. General public follows these newspapers to gauge each candidate with respect to these issues and builds up its opinion and inclination. Consider the following two sentences from *Chicago Tribune* and *LA Times* respectively:

What is clear from Christie's claims about NSA surveillance is that he 1) doesn't know what he's talking about, or 2) doesn't care.

Clinton dived deep into the challenges of being black in America and the structural racism embedded in the country's culture and economy.

We considered the 10 most important of these issues as aspects of these entities.

The dataset gathered in the first step was annotated for three things: entities, aspects and sentiments. Annotations were performed by a total of 11 annotators who were first assessed for linguistic proficiency and trained on our task. In an article, they annotated only those sentences that contained at least one entity and its aspect from the predefined inventories. They were given a list of all entities, aspects and multiple variations of an aspect present in the dataset. Sentiment annotations were performed on an interval scale ranging from -5 to +5, going from most negative to neutral at 0 and most positive. Each sentence was annotated by 2 annotators. For conformity of entities and aspects, we vetted all the sentences to contain the same set of entity aspect pairs using a third neutral annotator. We average the sentiment of both annotators for each entity-aspect pair to compensate for the errors due to subjectivity or bias. Further, we found very minor differences in conflicting sentiments for different entity-aspect pairs within a sentence. Hence for the sentiment prediction task, we extrapolate the average sentiment of the pairs to the whole sentence.

We measure the reliability of sentiment annotations using a strategy similar to the one used by (DanescuNiculescu-Mizil et al., 2013)[2]. For each batch of documents annotated by the same users, we calculate the pairwise correlation of the respective scores. We also collect the same data after random sampling from the distribution of sentiment scores. As opposed to the randomized scores, user annotated scores are highly correlated with each other with $p < 0.0001$ according to a paired t-test. See figure 1. To make sentiment prediction robust and to account for differing vantage points during annotation, we use 3 sentiment classes following the popular practice and convert the sentiment scores from 11-point interval into 3 discrete bins namely positive, negative and neutral classes. Sentences having scores $\in [3, 5]$ fall in the positive class, $[-3, -5]$ fall in negative class and the rest of the sentences fall in neutral class.

[1] Our dataset is available on request.

[2] Commonly used measures like Cohen's Kappa or Scott's Pi are not suitable for ordinal or interval data.

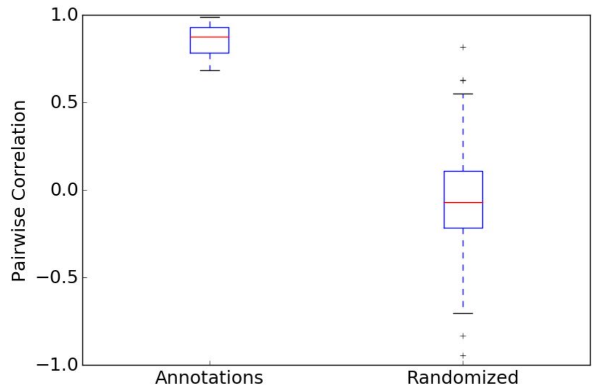

Figure 1: Inter-annotator pairwise correlation of sentiment annotations, compared with randomized annotations.

Our Algorithm

Our algorithm starts by establishing the presence of an entity of interest in a given document. After identification of all the entities, different aspects of those entities are extracted and entity-aspect relationship is established between them. Associating the correct entity with an aspect is important, especially as a sentence can mention a number of entities in close proximity. Finally, sentiment analysis is performed on the extracted entity-aspect pairs. Figure 2 illustrates our algorithm pipeline for a test document. Below, we briefly define the two essential components of these tasks.

Coreference Resolution: In large documents, human beings are mentioned directly as well as indirectly using pronouns. These implicit mentions known as coreferences, are matched with the entities they refer to using a technique called coreference resolution. We use the state-of-the-art coreference resolution system by (Lee et al., 2011) developed at Berkeley in our study.

Semantic Role Labeling (SRL): Semantic role labeling refers to the detection of semantic arguments associated with each verb in a sentence. For each target verb, words in a sentence serve as a semantic role of that verb. The following example makes understanding easier:

[$_{A0}$ He] [$_{AM\text{-}MOD}$ would] [$_{AM\text{-}NEG}$ n't] [$_{V}$accept] [$_{A1}$ anything of value] from [$_{A2}$ those he was writing about].

Here, the roles for the verb *accept* are defined in PropBank Frame scheme as *V* for verb, *A0* for acceptor, *A1* for thing accepted, *A2* for accepted-from, *AM-MOD* for modal and *AM-NEG* for negation.

Task 1: Mining Entity-Aspect Pairs

We use SENNA (Collobert et al., 2011) to extract different semantic roles from a sentence. For each sentence, we select the verbs labeled as *-V* and evaluate the corresponding roles for the words in the sentence. Entities from the sentence are extracted from the words found in *-A0*. *A0* generally contains the agent causing an action like *speaker*, *giver* and *keeper*. Most of the times, it refers to the main entity being discussed in the sentence. If an entity is not found in *A0*, it is searched in *A1* which can contain the target of an opinion as in the following:

NBC gave Donald Trump the ax Monday over remarks he made about Mexican immigrants in his presidential campaign kickoff.

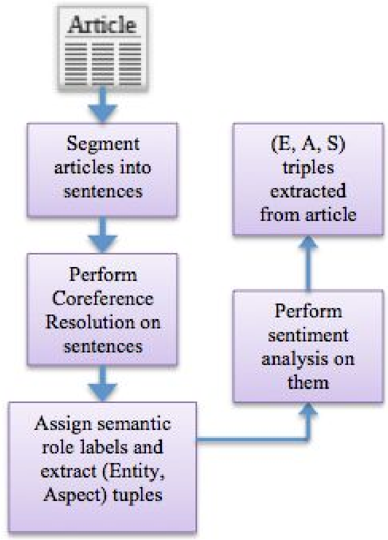

Figure 2: Algorithm pipeline for one document at test time. E, A and S denote entity, aspect and sentiment respectively.

After extracting the entities, we search for aspects in the rest of the roles in the sentence. Varying expressions of all aspects are found using manually constructed aspect maps. We analyzed our dataset to extract various patterns of expression of entity-aspect pairs as labeled by SRL, using which we associate aspects with their corresponding entities. We search for the aspects in semantic arguments labeled *-A1, -A2, -A3, -A4, -CAU, -TMP, -ADV* and *-LOC* with different conditions. *CAU* indicates causality of a given event and occurs occasionally in the explanation of an event in which the entity is involved. Similarly *LOC* indicates the location of a particular event in the presence of verbs like *live*. Names of countries like *Iraq* are found in *LOC* indicating the potential presence of aspects like foreign policy and terrorism. The label *TMP* refers to a temporal argument which signifies an aspect when mentioned with respect to a reference of time like *...when it comes to the economy*. Additionally, when an opinion is expressed using an adverb clause with verbs like *enacted* and *tried*, argument label *ADV* contains mentions of aspects.

Task 2: Sentiment Analysis

After the successful completion of Task 1, sentences are fed into the sentiment analysis module for sentiment prediction. We modify the Convolutional Neural Network (CNN) (Kim, 2014) for our 3-class problem using the extrapolated sentiment and leverage the distributed word vector representations in *word2vec* (Mikolov et al., 2013). The network architecture consisted of a simple CNN having one layer of convolution. The layer applies multiple non-linear filters to varying sized windows of words to produce feature maps. It then uses max-over time pooling over these features maps to select the most important features, which are then passed on to a fully connected softmax layer for prediction of probabilities for each of the sentiment classes. The network also uses static and non-static channels of word vectors to fine tune the initial *word2vec* feature vectors.

Evaluation

For both the tasks and baselines, we perform a 5-fold cross validation of our dataset. Due to similarity in prediction tasks and being state-of-the-art, we construct our baselines after slight modifications of baselines of the two subtasks of SemEval 2015 Task 12.

Task 1: For each sentence in our test set, we collect statistics for the correct number of entity-aspect pairs returned by our algorithm. We then calculate Precision, Recall and F1 measure for each sentence and use micro-averaging for overall evaluation. As a baseline, we train a support vector machine (SVM) with linear kernel using bag-of-words approach with tf-idf values of unigram features for sentences.

Task 2: Instead of treating classification results in a binary fashion, we calculate the distance of predicted label from its target. For the task at hand, we argue that predicting for instance neutral class for a positive sentence is less severe an error than predicting a negative class for it. Using the formula given below, we calculate sentiment score for one sentence *t* with improved accuracy over:

$$score(t) = \frac{|s_{predicted} - s_{target}|}{s_{max} - s_{min}}$$

$s_{predicted}$ is the predicted sentiment label, s_{target} is the target label, s_{max} and s_{min} are the largest and smallest possible sentiment values[3]. After calculating individual sentiment scores, they are averaged over the whole dataset for final accuracy. Our baseline for sentiment task is similar to the one developed for Slot3 of Task 12 in SemEval-2015. We train a support vector machine (SVM) with linear kernel. Each sentence is represented by a feature vector containing bag-of-words features and entity-aspect pairs present within.

Experimental Results

Method	P	R	F1
Automatic Coreference	62.52	64.92	63.94
Manual Coreference	64.20	77.31	70.14
Baseline	50.31	57.64	53.72

Table 1: Results of Task 1: Entity-Aspect pair extraction. P and R denote micro-averaged precision and recall. Manual coreference was done for a sample of sentences.

Results Task 1

On the extraction of entity-aspect pairs, we obtained an F1 of 64% as compared to 54% by our baseline. Analysis revealed that many sentences reported low recall due to error propagated forward from coreference resolution. In particular, its performance was negligible in resolving coreferring noun phrases like *the governor, both leaders*, etc. To verify, we manually resolved coreferences on a sample of sentences which gave an F1 score of 70%. Table 1 gives a summary of results.

Method	Score
Simple CNN	62.2
CNN non-static	63.5
Baseline	53.4

Table 2: Results of Task 2: Sentiment Analysis.

[3] s_{max} and s_{min} are 1 and -1 respectively.

Results Task 2

Our CNN model outperforms the baseline by about 10%. Apart from the static pre-trained word vectors by *word2vec*, we also experimented with fine tuning of word vectors by back propagating error derivatives to the input layer, resulting in slight improvement of results. Summary of the results is given in Table 2.

Application to US Presidential Elections 2016

While media's primary role in any democratic process is maintaining a high level of impartiality and integrity, cases of bias, slant, agenda setting etc. from news sources are often observed. Using the results from 2012 presidential elections as a proxy for political leaning of states, we analyzed the frequency and intensity of news coverage and found interesting evidence in our dataset of newspapers catering to a partisan audience by focusing on different aspects of the candidates in different states.

We began by looking for a systematic variation in different traits that newspapers focus on for each candidate. Figure 3 shows the coverage for top candidates in *Chicago Tribune* (CT), *LA Times* (LT) and *Houston Chronicle* (HC). Our analysis suggests that pro-democratic newspapers like LT and CT gave more coverage to race relations and economy compared to newspapers with a pro-republican endorsement pattern that focused more on health care, foreign policy and immigration. The data also suggested immigration as a major campaign issue for HC while economy featured more significantly for CT and LT. This could perhaps be best explained by newspapers catering to consumer preferences in these states. California is a very good similar example where the state has mostly voted Democratic based on strong support from minority voters since 1992. Trump has an all-negative coverage in LA while Clinton gets a mostly positive sentiment in all aspects.

Aspects	Before	After
Race Relations	12	88
Gender Issues	13	87
Basic Rights	36	64
Immigration	29	71
Taxes	60	40
Health Care	20	80
Foreign Policy	34	66
Security	14	86
Terrorism	42	58
Economy	29	71

Table 3: Trump Effect on campaign issues measured via %age of total mentions before and after Trump's declaration.

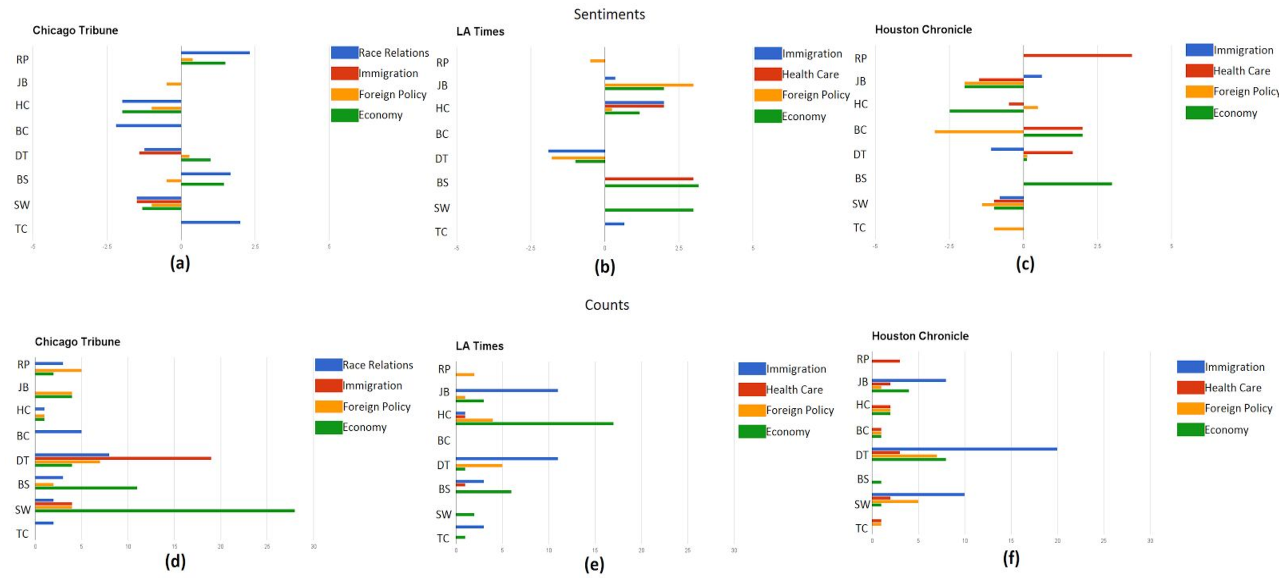

Figure 3: Aspect coverage and sentiment across three different newspapers. RP: Rand Paul, JB: Jeb Bush, HC:Hillary Clinton, BC: Ben Carson, DT: Donald Trump, BS: Bernie Sanders, SW: Scott Walker, TC: Ted Cruz

Looking for variations in the different aspects in Figure 3, we see clear evidence of differential focus in coverage across newspapers. CT has a negative sentiment for Clinton's race relations and economy while the overall coverage is positive on health care, foreign policy and gender issues. CT's coverage mostly focuses on race, gender issues, basic rights and immigration on Donald Trump. HC however has been traditionally pro-republican and operates in a more conservative market and hence looks positively at Trump's stance on taxes, health care, terrorism while criticizing Clinton on economy, security, healthcare and economic policies. Similar to this trend LT is firmly behind Clinton on economy, race, gender issues while being highly critical of Trump's race relations, immigration and foreign policies. Interestingly also, while LT and CT have the same parent companies, the coverage in the different newspaper clearly seems to be dictated by consumer preferences of the local population. CT prefers Bernie Sanders over Clinton across all aspects including basic rights and economy. However, LT makes no mention of these aspects when talking about him and leans heavily towards Clinton. HC on the other hand has very little or no mention of Bernie Sanders at all.

Next, we look for the evidence of the so-called *"Trump effect"* in our data corpus. Donald Trump announced his candidacy on the 16th of June 2015 and we partitioned our data to see how it affected the overall presidential race. Table 3 shows the mentions of each aspect before and after June 16th by any candidate in all newspapers as a percentage of total mentions. One of the significant things to note is the increased emphasis on race relations, gender issues, health care and security while important issues such as taxes and foreign policy received a lot less treatment overall. Correlating it with Figure 3, we can see that while immigration was always an issue in California and Texas, CT had no mention of immigration before Trump joined the race. In short, immigration has become the second most popular theme behind economy post June 16th. Another interesting effect is Jeb Bush's sentiment ratings increasing in pro-democratic newspapers like CT and LT where the average sentiment for him went from -.33 to .17 and .71 to 1.4 and decreasing in pro-republican strongholds such as HC and *Dallas News* where he went from between .83 and .89 to -1.45 and -2 respectively. This is perhaps in line with Trump playing to the audience in these states. Trump joining the race has also seen a steady positive sentiment for Rand Paul across all demographics.

In a close election, ABSA frameworks such as ours can be used in useful, interesting and compelling ways as a means of focusing campaign funds on targeted advertising campaigns in swing states based on real-data analysis.

Conclusion and Future Work

In this paper, we explore ABSA in the context of large documents for human entities. We construct a first-of-its-kind dataset containing US news articles. We then present and discuss a novel algorithm for extraction and analysis of entity-aspect pairs from these documents. The algorithm employs coreference resolution, semantic role labeling and sentiment classification. Our results demonstrate significant improvement of our algorithm over baseline methods for both the tasks of entity-aspect pair extraction and sentiment analysis. In future, ABSA in large documents should be performed beyond sentence level to capture more complex semantics and broader context. Moreover, specific parts of the sentence should be isolated to perform sentiment analysis for individual entity-aspect pairs. Finally, coreference resolution needs improvement particularly for pronomials to identify the entities in complex textual settings.

REFERENCES

Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G. A., and Reynar, J. 2008. “Building a sentiment summarizer for local service reviews,” *In WWW Workshop on NLP in the Information Explosion Era - volume 14.*

Blei, D. M., Ng, A. Y., and Jordan, M. I. 2003. “Latent dirichlet allocation,” *The Journal of machine Learning research* (3), pp. 993–1022.

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., and Kuksa, P. 2011. “Natural language processing (almost) from scratch,” *The Journal of Machine Learning Research* (12), pp. 2493– 2537.

Cui, H., Mittal, V., and Datar, M. 2006. “Comparative experiments on sentiment classification for online product reviews,” *In AAAI - volume 6*, pp. 1265–1270.

Danescu-Niculescu-Mizil, C., Sudhof, M., Jurafsky, D., Leskovec, J., and Potts, C. 2013. “A computational approach to politeness with application to social factors,” arXiv preprint arXiv:1306.6078.

Jiang, L., Yu, M., Zhou, M., Liu, X., and Zhao, T. 2011. “Target-dependent twitter sentiment classification,” *In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, Association for Computational Linguistics, pp. 151– 160.

Jiang, P., Zhang, C., Fu, H., Niu, Z., and Yang, Q. 2010. “An approach based on tree kernels for opinion mining of online product reviews,” *In Data Mining (ICDM)*, 2010 IEEE 10th International Conference, pp. 256–265. IEEE.

Kennedy, A., and Inkpen, D. 2006. “Sentiment classification of movie reviews using contextual valence shifters,” Computational intelligence (22:2), pp. 110– 125.

Kessler, J. S., and Nicolov, N. 2009. “Targeting sentiment expressions through supervised ranking of linguistic configurations,” *In ICWSM.*

Kim, S., and Hovy, E. 2006. “Extracting opinions, opinion holders, and topics expressed in online news media text,” *In Proceedings of the Workshop on Sentiment and Subjectivity in Text*, Association for Computational Linguistics, pp. 1–8.

Kim, Y. 2014. “Convolutional neural networks for sentence classification,” arXiv preprint arXiv:1408.5882.

Lee, H., Peirsman, Y., Chang, A., Chambers, N., Surdeanu, M., and Jurafsky, D. 2011. “Stanford’s multi-pass sieve coreference resolution system at the conll-2011 shared task,” *In Proceedings of the Fifteenth Conference on Computational Natural Language Learning: Shared Task*, Association for Computational Linguistics, pp 28–34.

Liu, P., Joty, S., and Meng, H. 2015. “Fine grained opinion mining with recurrent neural networks and word embeddings,” *In Conference on Empirical Methods in Natural Language Processing (EMNLP 2015).*

Liu, Q., Liu, B., Zhang, Y., Kim, D. S., and Gao, Z. 2016. “Improving Opinion Aspect Extraction using Semantic Similarity and Aspect Associations,” *In Proceedings of Thirtieth AAAI Conference on*

Artificial Intelligence, Phoenix, Arizona, USA.

Long, C., Zhang, J., and Zhut, X. 2010. "A review selection approach for accurate feature rating estimation," *In Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, Association for Computational Linguistics, pp. 766–774.

McAuley, J., Leskovec, J., and Jurafsky, D. 2012. "Learning attitudes and attributes from multi-aspect reviews," *In Data Mining (ICDM)*, 2012 IEEE 12th International Conference, pp. 1020–1025. IEEE.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. 2013. "Efficient estimation of word representations in vector space," arXiv preprint arXiv:1301.3781.

Moghaddam, S., and Ester, M. 2012. "Aspect based opinion mining from online reviews," *In Tutorial at SIGIR Conference*.

Nakov, P., Ritter, A., Rosenthal, S., Sebastiani, F., and Stoyanov, V. 2014. "SemEval-2016 Task 4: Aspect based sentiment analysis" *In Proceedings of the 10th international workshop on semantic evaluation (SemEval 2016)*, pp. 1–18.

Pang, B., and Lee, L. 2008. "Opinion mining and sentiment analysis," Foundations and trends in information retrieval (2:1-2), pp. 1–135.

Peters, G., and Woolley, J. T. 2012. "2012 General election editorial endorsements by major newspapers," http://www.presidency.ucsb.edu/data/ 2012_newspaper_endorsements.php.

Pontiki, M., Galanis, D., Papageogiou, H., Manandhar, S., and Androutsopoulos, I. 2015. "Semeval-2015 task 12: Aspect based sentiment analysis," *In Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*, Denver, Colorado.

Pontiki, M., Papageorgiou, H., Galanis, D., Androutsopoulos, I., Pavlopoulos, J., and Manandhar, S. 2014. "Semeval-2014 task 4: Aspect based sentiment analysis," *In Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014)*, pp. 27–35.

Poria, S., Cambria, E., and Gelbukh, A. 2016. "Aspect Extraction for opinion mining with a deep convolutional neural network," *In Knowledge-Based Systems - volume 108,* pp. 42-49

Prasojo, R. E., Kacimi, M., and Nutt, W. 2015. "Entity and aspect extraction for organizing news comments," *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, ACM.

Rosenthal S., Farra N. and Nakov P., 2017. "SemEval-2017 Task 4: Sentiment Analysis in Twitter" *In Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval '17),* Vancouver, Canada.

Sauper, C., Haghighi, A., and Barzilay, R. 2011. "Content models with attitude," *In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, Association for Computational Linguistics, pp. 350–358.

Wang, W., Pan, S. J., Dahlmeier, D., and Xiao., X. 2016, "Recursive Neural Conditional Random Fields for Aspect-based Sentiment Analysis," *In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing,* Austin, Texas.

Wang, Y., Huang, M., and Zhu, X. 2016. "Attention-Based LSTM for Aspect-level Sentiment Classification," *In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, Texas.

Zhang, L., and Liu, B. 2014. "Aspect and entity extraction for opinion mining," *In Data mining and knowledge discovery for big data*, Springer, pp. 1–40.