Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM 2018)

Opinion Context Extraction for Aspect Sentiment Analysis

Anil Bandhakavi,[1] Nirmalie Wiratunga,[1] Stewart Massie,[1] Rushi Luhar[2]
[1]School of Computing, Robert Gordon University, Aberdeen, UK
a.s.bandhkavai, n.wiratunga, s.massie@rgu.ac.uk
[2]SentiSum, London, UK rushi@sentisum.com

Abstract

Sentiment analysis is the computational study of opinionated text and is becoming increasing important to online commercial applications. However, the majority of current approaches determine sentiment by attempting to detect the overall polarity of a sentence, paragraph, or text window, but without any knowledge about the entities mentioned (e.g. restaurant) and their aspects (e.g. price). Aspect-level sentiment analysis of customer feedback data when done accurately can be leveraged to understand strong and weak performance points of businesses and services; and can also support formulation of critical action steps to improve performance. In this paper we focus on aspect-level sentiment classification, studying the role of opinion context extraction for a given aspect and the extent to which traditional and neural sentiment classifiers benefit when trained using the opinion context text. We propose four methods to aspect context extraction using lexical, syntactic and sentiment co-occurrence knowledge. Further, we evaluate the usefulness of the opinion contexts for aspect-sentiment analysis. Our experiments on benchmark data sets from SemEval and a real-world dataset from the insurance domain suggests that extracting the right opinion context is effective in improving classification performance. Specifically combining syntactical features with sentiment co-occurrence knowledge leads to the best aspect-sentiment classification performance.

1 Introduction

Sentiment Analysis (SA) is critical for an increasing number of applications and industries, due to the vast amounts of opinionated content generated on social media about the products and services they offer. Recent statistics suggests that on average in a given minute, over 400,000 Twitter posts are shared, about 300,000 Facebook statuses updated, about 25,000 items purchased from Amazon[1].

Sentence level analysis of opinionated content is common but these ignore sentence structure and semantic constructs (Mohammad, Kiritchenko, and Zhu 2013; Joulin et al. 2016). Increasingly, more granular analysis is needed to better understand the target of the opinion, referred to as the aspect, as well as the context within which that sentiment is being expressed (Laddha and Mukherjee 2016). Indeed the ability to analyze opinionated content beyond just the surface level is crucial to discover meaningful business insights for companies. For instance given *the food was amazing but the service could have been better*, we can observe that although the overall sentence polarity can be viewed as being positive, there is to some degree a level of negative polarity also being expressed towards aspect, *service*, when sentence context, *the service could have been better*, is inspected more closely.

In this paper our focus is on aspect context extraction also called Opinion Context Extraction (OCE). Accordingly, we present results from a comparative study of alternative OCE methods for aspect sentiment analysis using two SemEval datasets (Pontiki et al. 2016) and a real-world business data set of customer reviews for the insurance domain. Specifically, we introduce a syntactic windowing method to OCE, which unlike the popular lexical windowing approaches (Bing 2012), exploits sentence structure and dependencies to extract the sub components of a sentence that are considered most relevant to an aspect for sentiment computation. Further, we also introduce an OCE method that combines knowledge of the syntactic dependencies in text with sentiment co-occurrence statistics between aspects and sentiment words found in a high coverage sentiment lexicon. Whilst both supervised and unsupervised approaches to SA stand to benefit from OCE, in this paper we focus on supervised models (Kim 2014; Socher et al. 2013; Kiritchenko et al. 2014), whereby the extracted context of an aspect, is used by a classifier to assign a sentiment polarity label for the aspect.

2 Related Work

2.1 Aspect Extraction

One approach is to extract all the different nouns and noun phrases from the text and consider them as candidate aspect terms (Hu and Liu 2004). Schouten develop a co-occurrence based method for category discovery using a dictionary-based sentiment classification algorithm through which aspects can be identified by an annotation process (Schouten and de Jong. 2014). Alternatively, aspect extraction can be modeled as a sequential labeling task with features extracted for CRF training (Zhiqiang and Wenting. 2014), (Malhotra

[1]http://www.visualcapitalist.com/happens-internet-minute-2017/

et al. 2014), (Brychcn and Steinberger. 2014).

In this work we extract aspects using the knowledge of domain experts and focus on evaluating different approaches for extracting the context associated with each aspect and also the impact of such contexts on different sentiment classifiers to predict aspect level sentiment.

2.2 Opinion Context Extraction

Subjective expression extraction has traditionally used sequence models, such as CRFs (Choi et al. 2005). An alternative approach is to employ a Hidden Markov Model (HMM) over words, and model the latent topics as states in the HMM to discover the product properties (often aspects) and the associated attributes (such as pos/neg polarities) separately (Sauper, Haghighi, and Barzilay. 2011). Yang and Cardie. (2012, 2014) use a semi-CRF based approach which allows sequence labeling at segment level and (Yang and Cardie. 2014) is employed for opinion expression intensity and polarity classification. However, the above works focus on generic subjective expressions as opposed to aspect specific opinion-sentiment phrases. Our work also takes advantage of dependency parsers and sentiment co-occurrence statistics to extract candidate opinion phrases for aspect-sentiment analysis.

2.3 Sentiment Analysis

The state-of-the-art in sentiment analysis includes diverse techniques, such as rule-bases, lexicons (Alec Go and Huang 2009), machine learning (Mohammad, Kiritchenko, and Zhu 2013; Nakagawa, Inui, and Kurohashi 2010; Arora et al. 2010) and deep learning (Ribeiro et al. 2016), (Joulin et al. 2016), (Kim 2014; Socher et al. 2013).

In this work we do not propose a new sentiment classifier, however we evaluate the effectiveness of different state-of-the-art sentiment analyzers trained using the text generated from the proposed opinion context extraction methods for predicting aspect-level sentiment.

3 Opinion Context Extraction Approaches

In this section we formalize the different opinion context extraction approaches.

3.1 Sentence level context

The baseline strategy for context extraction is to simply use the entire sentence as containing the relevant context for any given aspect, a, in that sentence. Sentiment classification is applied to the entire sentence bearing a and the corresponding prediction assigned to a:

$$sentiment\_classifier(a,\text{sentence}(a)) \tag{1}$$

Where $sentiment\_classifier()$ is a function that predicts the sentiment expressed, in relation to the aspect as positive, negative or neutral. This context extraction approach is reasonable if the sentence contains only a single aspect and the sentiment words in the sentence are used to express opinion towards that aspect. However it is less effective when the sentence contains multiple aspects with contrasting sentiment expressed towards each of them.

3.2 Lexical window of context

In this approach we identify a window of k words around an aspect as the context window from which to extract text for sentiment analysis. The size of the window is chosen empirically to be 3. More formally, let a sentence be denoted as $S = \{w_1, \ldots, w_n\}$. Assuming $w_x \in S$ as the aspect a, the lexical window of context for w_x is extracted as follows:

$$Context_{lex}(w_x) = \begin{cases} [w_{x-k} : w_{x+k}] & if x<n-k \text{ and } x>k \\ [w_1 : w_{x+k}] & if x<k \\ [w_{x-k} : w_n] & if x+k>n \end{cases} \tag{2}$$

This approach assumes that the opinion words targeting an aspect occur close by, in the window of k words from the aspect, and that extracting the words within that window gives a useful bag-of-words for analyzing the sentiment of the aspect.

Sentiment classification is applied using the lexical context associated with the aspect a and the corresponding prediction is assigned to a as follows:

$$sentiment\_classifier(a,Context_{lex}(a)) \tag{3}$$

3.3 Syntactical window of context

With complex sentences involving multiple aspects one cannot rely solely on adjacency of text as a cue to context identification.

In the syntactically-informed windowing approach, we study the dependency relationships within a sentence to extract the window of k words to form an aspect's context. Unlike the lexical window which ignores the syntactic relationships between words, this approach incrementally traverses the dependency parse tree, starting from the aspect (node) in either direction to arrive at the context text for sentiment analysis. The standard tool used in natural language processing for learning the syntactic structure of sentences is a dependency parser. In this work we use trees constructed through Spacy[2] to extract the relevant text window.

More formally, let a sentence be denoted as $S = \{w_1, \ldots, w_n\}$. Let $T = \{t_1, \ldots, t_n\}$ be the dependency tree corresponding to S, where each $t_i \in T$ is a triplet $(w_i, parent(w_i), children(w_i))$ where $parent(w_i) \in S$ and $children(w_i) \in S$.

Once the context text is extracted sentiment classification is applied using the discovered syntatic context associated with the aspect a and the corresponding prediction is assigned to a as follows:

$$sentiment\_classifier(a,Context_{synt}(a)) \tag{4}$$

3.4 Syntactical sentiment weighted co-occurrence window of context

A sentiment-rich corpus of text can be used to learn how often a list of sentiment words and aspects co-occur. Furthermore, this knowledge can be used to guide the traversal of the dependency tree to collect the words that influence the aspect unlike the previous approach which uses distance between words within the tree. Unlike 3.3, here we are able

[2]https://spacy.io/

to commit to the most promising sub-tree thereby disregarding neighboring sub-trees that are less promising in terms of aspect sentiment relatedness.

More formally, let a sentence be denoted as $S = \{w_1, \ldots, w_n\}$. Let T be the dependency tree corresponding to S and T^* be a subtree of T. Let A be the set of aspects identified for a corpus of reviews. Sentiment classification is applied using the discovered context associated with the aspect a and the corresponding prediction is assigned to a as follows:

$$sentiment\_classifier(a, Context_{synt\_senti}(a)) \quad (5)$$

4 Sentiment Classifier

In this section we briefly describe the sentiment classifiers used in this work for evaluating the quality of the opinion context extraction approaches for aspect-level sentiment analysis. We selected a diverse set of sentiment classifiers ranging from feature engineering-based (e.g. NRC sentiment) (Kiritchenko et al. 2014) to shallow neural networks (e.g. fastText) (Joulin et al. 2016) to deep neural networks (e.g.convolutional neural network (CNN)) (Kim 2014).

5 Evaluation

The aim of the evaluation is to validate the usefulness of the proposed opinion context extraction methods for effective aspect-sentiment analysis. Our evaluation is a comparative study of the performance of the different opinion context extraction methods using aspect sentiment analysis.

5.1 Datasets

We used three different data sets (customer reviews) from the domains of restaurants and insurance for our evaluation. The restaurant data sets are official benchmark data sets from the SemEval competition for 2015[3] and 2016[4]. The data set for the insurance domain is a commercial data set.

5.2 Results and Analysis

In this section we present the results obtained for different opinion context extraction approaches in aspect-sentiment classification task.

Aspect sentiment classification Here we use sentiment classification as a means to find out how effective each extraction method is for aspect-level sentiment prediction. Table 1 shows the aspect level sentiment prediction results (best results highlighted in bold) for SemEval and the insurance data sets. It was observed that using sentence text as the context for aspect sentiment analysis is a strong competition for the other methods which use only a span of text within the sentence as context text for aspect sentiment prediction. We believe this is due to the presence of single aspect bearing sentences, where the entire sentence is a description about one aspect and its sentiment. However context-aware methods outperform the sentence based method suggesting that extracting the right window of text around the aspect is useful for aspect sentiment classification. Further amongst the opinion context methods in general the lexical context based approach has the weakest performance. This could be due to the ineffectiveness of the lexical window in capturing the relevant sentiment words that target the aspects.

Table 1: Aspect Sentiment Analysis Results on SemEval-2015, SemEval-2016 and Insurance Data

	Sentence	Lexical	Synt (k=4)	Synt_Senti
FastText on SemEval-2015 data				
F_score	61.98	57.90	66.21	67.32
Accuracy	65.28	58.45	68.24	69.45
NRC Sentiment on SemEval-2015 data				
F_score	68.42	69.78	66.56	**72.34**
Accuracy	69.76	72.81	69.37	**75.24**
CNN on SemEval-2015 data				
F_score	63.23	59.56	68.34	70.38
Accuracy	67.14	60.45	70.42	72.26
FastText on SemEval-2016 data				
F_score	72.69	64.73	76.33	77.85
Accuracy	73.29	62.72	76.44	77.15
NRC Sentiment on SemEval-2016 data				
F_score	74.36	77.12	**78.55**	78.12
Accuracy	74.68	77.83	**79.72**	78.67
CNN on SemEval-2016 data				
F_score	72.14	65.68	75.26	76.12
Accuracy	72.86	63.24	75.89	76.68
FastText on Insurance data				
F_score	74.42	61.96	70.44	77.35
Accuracy	75.72	62.85	70.09	77.45
NRC Sentiment on Insurance data				
F_score	77.56	78.34	80.46	**80.78**
Accuracy	78.74	80.05	82.73	**82.87**
CNN on Insurance data				
F_score	75.67	63.00	73.49	77.48
Accuracy	76.12	64.12	74.57	78.67

For the syntactic context approach the best performance of sentiment analysis is when value of k is 4. This suggests that having longer context which include immediate as well as distant syntactic dependents for the aspect is more effective to capture the relevant opinion words that target the aspects thereby boosting the sentiment classifier performance. Further the approach which combines the syntactic dependency information with the sentiment co-occurrence information to extract the opinion context surrounding an aspect either records the best performance (SemEval-2015, insurance) or is comparable with the performance of the approach which uses only the syntactic context (SemEval-2016). Overall these results suggests that the sentiment co-occurrence guided sub-tree selection heuristic for disambiguating aspect contexts, is specifically beneficial for multi-aspect sentence analysis.

Amongst the different sentiment learners used, NRC sentiment classifier performs best consistently outperforming its neural counterparts. Amongst the neural methods CNN performs better than fastText. This suggests that with more depth in the neural network there is scope for learning better predictive models. We believe that NRC sentiment classifier

[3] http://alt.qcri.org/semeval2015/task12/

[4] http://alt.qcri.org/semeval2016/task5/

learns features that are complementary and are collectively effective in predicting the sentiment at the aspect level. Further in the case of NRC since it uses both the sentence level text and opinion context within the sentence for feature extraction we found it to be having an advantage over the neural classifiers which consider only the opinion context text as input.

6 Conclusions

In this paper we investigate the role of opinion context extraction for aspect-level sentiment classification with an aim to evaluate the extent to which traditional and neural sentiment classifiers benefit when trained using the opinion context text. We propose four methods to extract opinion contexts surrounding aspects using lexical, syntactic and sentiment co-occurrence knowledge through standard aspect-sentiment classification tasks. Our experiments on benchmark data sets from SemEval and a real-world dataset from the insurance domain suggests that extracting the right opinion context is effective in improving classification performance. Specifically our heuristic which combines syntactical features with sentiment co-occurrence knowledge leads to the best aspect-sentiment classification performance.

References

Alec Go, R. B., and Huang, L. 2009. Twitter sentiment classification using distant supervision. *Processing*.

Arora, S.; Mayfield, E.; Penstein-Rosé, C.; and Nyberg, E. 2010. Sentiment classification using automatically extracted subgraph features. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text pp 131–139*.

Bing. 2012. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.

Brychcn, Tom, M. K., and Steinberger., J. 2014. ”uwb: Machine learning approach to aspect-based sentiment analysis.” semeval 2014. In *SemEval (2014): 817*.

Choi, Y.; Cardie, C.; Riloff, E.; and Patwardhan., S. 2005. Identifying sources of opinions with conditional random fields and extraction patterns. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing, pp 355-362.*

Hu, M., and Liu, B. 2004. Mining opinion features in customer reviews. In *Proceedings of the 19th National Conference on Artifical Intelligence (AAAI 2004)*.

Joulin, A.; Grave, E.; Bojanowski, P.; and Mikolov, T. 2016. Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.

Kim, Y. 2014. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp 1746-1751, October 25-29, 2014*.

Kiritchenko, S.; Zhu, X.-D.; Cherry, C.; and Mohammad, S. 2014. Nrc-canada-2014: Detecting aspects and sentiment in customer reviews. In *SemEval@COLING*.

Laddha, A., and Mukherjee, A. 2016. Extracting aspect specific opinion expressions. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp 627-637,*.

Malhotra; Nishtha; Malhotra; and Nishtha. 2014. ”sap-ri: A constrained and supervised approach for aspect-based sentiment analysis.”. In *SemEval 2014 : 517*.

Mohammad, S. M.; Kiritchenko, S.; and Zhu, X. 2013. Nrc-canada: Building the state-of-the-art in sentiment analysis of tweets. In *Seventh International Workshop on Semantic Evaluation (SemEval 2013), pp 321–327*.

Nakagawa, T.; Inui, K.; and Kurohashi, S. 2010. Dependency tree-based sentiment classification using crfs with hidden variables. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics pp 786–794*.

Pontiki, M.; Galanis, D.; Papageorgiou, H.; and Androutsopoulos, I. 2016. Semeval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of SemEval-2016, pages 1930, San Diego, California, June 16-17, 2016, Association for Computational Linguistics*.

Ribeiro, F. N.; Araujo, M.; Goncalves, P.; and Marcos Andre Goncalves, a. F. B. 2016. Sentibench- a benchmark comparision of state-of-the-paractice sentiment analysis methods. *Springer EPJ Data Science*.

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, pp 350-358*.

Schouten, Kim, F. F., and de Jong., F. 2014. Commit-p1wp3: A co-occurrence based approach to aspect-level sentiment analysis. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014).*

Socher, R.; Perelygin, A.; Wu, J.; Chuang, J.; Manning, C.; Ng, A.; and Potts, C. 2013. Recursive deep models for semantic compositionality over a seniment treebank. In *Proc of EMNLP 2013*.

Yang, B., and Cardie., C. 2012. Extracting opinion expressions with semi-markov conditional random fields. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, pp 1335-1345.*

Yang, B., and Cardie., C. 2014. Joint modeling of opinion expression extraction and attribute classification. *Transactions of the Association for Computational Linguistics, 2:505516*.

Zhiqiang, T., and Wenting., W. 2014. Dlirec: Aspect term extraction and term polarity classification system (2014).