A Machine Learning approach for Subjectivity Classification based on Positional and Discourse Features

Jose M. Chenlo and David E. Losada

Centro de Investigación en Tecnoloxías da Información (CITIUS)
Universidad de Santiago de Compostela, Spain
{josemanuel.gonzalez,david.losada}@usc.es
http://www.gsi.dec.usc.es/ir

Abstract. In recent years, several machine learning methods have been proposed to detect subjective (opinionated) expressions within on-line documents. This task is important in many Opinion Mining and Sentiment Analysis applications. However, the opinion extraction process is often done with rough content-based features. In this paper, we study the role of structural features to guide sentence-level subjectivity classification. More specifically, we combine classical n-grams features with novel features defined from positional information and from the discourse structure of the sentences. Our experiments show that these new features are beneficial in the classification of subjective sentences.

Key words: Information Retrieval, Opinion Mining, Subjectivity Classification, Sentiment Analysis, Machine Learning, Rhetorical Structure Theory

1 Introduction

With the advent of the social web, opinions have become a key component in many on-line repositories [1]. These new opinion-oriented resources demand advanced classification processes able to skim off the opinionated texts to reveal the subjective parts. Extracting opinions from text is challenging and poses many problems that cannot be merely solved with lexicon-based approaches. These difficulties are caused by the subjectivity of a document being not so much conveyed by the sentiment-carrying words that people use, but rather by the way in which these words are used. We argue that the study of sentence positional information and intra-sentence discourse structure can help to tackle this issue. For instance, people tend to summarise their viewpoints at the end of the text. Moreover, the rhetorical roles of text segments can effectively guide the opinion detection process. For example, the sentence *"Nevertheless it is undeniable that economic disparity is an important factor in this ethnic conflict"* contains an attribution relationship between the nucleus of the sentence (*"that economic disparity is an important factor in this ethnic conflict"*) and its satellite (*"Nevertheless it is undeniable"*). The presence of this relation helps to understand that the writer is expressing his/her point of view (satellite) about the statement presented in the nucleus. This type of rhetorical clue is potentially valuable to detect opinions.

In this paper we combine bag of words features, such as unigrams or bigrams, with features computed from sentiment lexicons and with more advanced positional and rhetorical features. Our results show that the combined use of these features leads to an accurate classification of subjective sentences. To the best of our knowledge, this is the first attempt to combine rhetorical, content-based and positional features for a fine-grained (i.e., sentence-level) estimation of subjectivity. As argued in our related work section, other studies have explored the role of rhetorical features in Opinion Mining (OM) but previous efforts are mostly based on coarse-grained tasks (e.g., categorising the overall orientation of a movie review). Document-level sentiment classification is too crude for most applications and we need to move to the sentence level to support a more advanced analysis of sentiments [2].

2 Sentence Features

We focus on a two-class (subjective vs. non-subjective) classification of sentences[1] and take into account the following traditional and advanced features to build our classifiers:

Unigram & Bigram features. These are binary features based on the appearance of unigrams and bigrams in the sentence[2].

Sentiment Lexicon features. These features are based on counting the sentiment-bearing terms that occur in the sentence. The sentiment lexicon was obtained from OpinionFinder [3], which is a well-known subjectivity classifier. We include the number and percentage of opinionated terms in a sentence as features for our classifiers. We also consider the number and percentage of interrogations and exclamations in the sentence. The ability of these features to detect opinions has been demonstrated in other studies [4].

Rhetorical features. Rhetorical Structure Theory (RST) [5] is one of the leading discourse theories. This theory explains how texts can be split into segments that are rhetorically related to one another. Within this structure, text segments can be either nuclei or satellites, with nuclei being assumed to be more significant than satellites with respect to understanding and interpreting a text. Many types of relations between text segments exist; the main paper on RST defines 23 types of relations [5]. A satellite may for instance be an elaboration, an explanation or an evaluation on what is explained in a nucleus. We used

[1] A subjective sentence may not express any sentiment ("I think that he left") and an objective sentence can imply an opinion ("The phone broke in three days") [2]. We are interested in detecting opinions at sentence-level and, therefore, we search for sentences expressing either explicit or implicit sentiments. However, we use subjective and opinionated interchangeably to refer to sentences that implicitly or explicitly express opinions.

[2] Unigrams and bigrams with less than 4 occurrences in the collection are removed.

SPADE (Sentence-level PArsing of DiscoursE) [6], which creates RST trees for individual sentences and we include binary features associated to the appearance of every type of RST relation in a given sentence. Observe that we make an intra-sentence RST analysis. The study of inter-sentence RST analysis is an interesting challenge that is out of the scope of this paper. The relation types considered are reported in Table 1.

Length features. These features encode the length of the sentence, the length of the nucleus and the length of the satellite of the sentence (all of them computed as the total number of words). These features could be indicative of the way in which people write opinions. For instance, a factual sentence could be longer than an opinionated one.

Positional features. We encode the absolute position of the sentence within the document (e.g., 2 for the second sentence in the document) and its relative position (the absolute position normalised by the number of sentences in a document). We also include the number of sentences in the document as a feature. In this way we can represent if the sentence comes from a shrot or from a long text. These positional features could be highly indicative of opinions. For instance, writers tend to express their thoughts about the topic of the document at the end.

Table 2 summarises the sentence features considered in our study.

3 Experimental Design

To test our methods we need a collection of labelled sentences. We work here with the English version of the NTCIR-7 English MOAT Research collection [7]. This collection contains news from different sources and provides topics[3], documents that were assessed as relevant to the topics, and annotated data at sentence level. The annotations include both relevance and subjectivity information about the sentences, as well as the identification of the opinion holders. The labels were produced by three different assessors. We constructed our ground truth for subjectivity classification using a majority rule: a sentence is classified as subjective (resp. objective) if at least two assessors tagged it as subjective (resp. objective). As a consequence of this process, our ground truth is composed of 3584 sentences: 2697 sentences judged as objective and 887 judged as subjective. The preprocessing of this collection[4] led to a set of 2218 unigrams and 2812 bigrams.

Baseline. We measured the relative merits of our classification approach against a competitive baseline, OpinionFinder (OF) [3]. OF is a robust subjectivity classifier [8] that is powered by Naive Bayes classifiers trained using linguistic features. Basically, it uses linguistic patterns that are correlated with objectivity (resp. subjectivity) and then using them as features in a machine learning

[3] Textual representations of user needs. The information provided include title and narrative statements.

[4] We did not apply stemming and we did not remove stop words.

Table 1. RST relation types taken into account.

Relation	Description
attribution	Clauses containing reporting verbs or cognitive predicates related to reported messages presented in nuclei.
background	Information helping a reader to sufficiently comprehend matters presented in nuclei.
cause	An event leading to a result presented in nuclei.
comparison	Clauses presenting matters which are examined along with matters presented in nuclei in order to establish similarities and dissimilarities.
condition	Hypothetical, future, or otherwise unrealized situations, the realization of which influences the realization of nucleus matters.
contrast	Situations juxtaposed to situations in nuclei, where juxtaposed situations are considered as the same in many respects, yet differing in a few respects, and compared with respect to one or more differences.
elaboration	Rhetorical elements containing additional detail about matters presented in nuclei.
enablement	Rhetorical elements containing information increasing a readers' potential ability of performing actions presented in nuclei.
evaluation	An evaluative comment about the situation presented in the associated nucleus.
explanation	Justifications or reasons for situations presented in nuclei.
joint	No specific relation is assumed to hold with the matters presented in the associated nucleus.
temporal	Clauses describing events with a specific ordering in time with respect to events described in nuclei.

algorithm. Extraction patterns were created by applying a set of syntactic templates to the corpus. These patterns reflect syntactic relationships identified by a shallow parser [9]. Two classifiers are supported by OF: an accuracy classifier and a precision classifier. The first one yields the highest overall accuracy. It tags each sentence as either subjective or objective. The second classifier optimises precision at the expense of recall. We used the first classifier and we adopted the $F1$ score (computed with respect to the subjective class) to evaluate opinion detection effectiveness.

Classification method. In our experiments we used *liblinear* [10], which is a highly effective library for large-scale linear classification. This library handles Support Vector Machines (SVMs) classification and Logistic Regression classification with different regularisation and loss functions. These classifiers have shown to be very effective in many scenarios. We extensively tested all the classifiers supported by *liblinear* against the training collection to understand what classifier performs the best.

Table 2. List of sentence features.

Group	Feature
vocabulary	unigrams and bigrams.
Opinion	number of opinionated terms (pos. & neg.). number of exclamations and interrogations. number of opinionated terms (pos. & neg.) normalized by the length of the sentence. number of exclamations and interrogations normalized by the length of the sentence.
RST	Contains a satellite (binary feature). Contains a specific type of satellite: For each relation type reported in Table 1, we represent whether or not the sentence contains that relation type (one binary feature for each type of RST relation).
Position	Number of sentences in the document. Relative position of the sentence in the document. Absolute position of the sentence in the document.
Length	Length of the sentence. Length of the nucleus. Length of the satellite.

Training & Test. We randomly split the dataset into a training and test set, consisting of 75% and 25% of the sentences, respectively[5]. With the training set we applied 5-fold cross validation to set all the parameters of the classifiers and also to select the best performing classifier[6]. Observe that this two-class classification process is unbalanced: only 887 out of 3584 sentences are labelled as subjective. When dealing with unbalanced problems, discriminative algorithms such as SVMs or Linear Regression, which maximise classification accuracy, result in trivial classifiers that completely ignore the minority class [11]. Some of the typical methods to deal with this problem include oversampling the minority class (by repeating minority examples), under-sampling the majority class (by removing some examples from the majority class), or adjusting the misclassification costs. Oversampling the minority class results in considerable computational costs during training because it significantly increases the size of the training collection. Under-sampling the majority class is not an option for our problem because we have a small number of positive examples and we would need to remove most of the negative examples in order to have sets of positive

[5] We repeated this process 10 times and we averaged out the performance achieved to obtain a reliable estimation of effectiveness.

[6] Usually, the best classifier was a Logistic Regression classifier. The optimal value of the misclassification cost of the subjectivity class was often around 10. The optimal values for C were in the range (0,100].

examples and negative examples that are comparable in size. This massive removal of negative examples would result in much information being missed. We therefore opted for adjusting the misclassification costs to penalise the error of classifying a positive example as negative (i.e., subjective sentence classified as a non-subjective). The training process was designed to maximize the $F1$ score computed with respect to the subjective class. Next, we used the test set to evaluate the best performing classifier against unseen data.

4 Results

The test results are reported in Table 3. We experimented with different sets of features combined with unigrams and unigrams/bigrams representations. We also include OF's results against the same collection.

Table 3. Experimental results against the test dataset. The best performance for each column and for each set of features is bolded.

	Precision	Recall	F1
OpinionFinder	.4420	.4126	.4268
unigrams	.4926	.3855	.4325
+ Rhetorical	.4903	.4140	.4489
+ Positional	.4716	**.5033**	**.4869**
+ Length	.4571	.4846	.4704
+ Sentiment Lexicon	**.5077**	.4513	.4778
+ All	.4892	.4822	.4857
unigrams & bigrams	.5410	.3591	.4317
+ Rhetorical	.4903	.3576	.4248
+ Positional	.5045	.4573	.4797
+ Length	.4806	.4464	.4629
+ Sentiment Lexicon	**.5517**	.3883	.4558
+ All	.4858	**.5150**	**.5000**

Most of our methods outperform OF, being the unigrams/bigrams representation combined with all other features the best method ($F1$ score of 0.5). Analysing individually the sets of features we can observe that positional features seem to be highly indicative of subjectivity. The potential benefits of positional information to detect subjective information has been also shown in polarity estimation in blogs [12]. Features based on sentiment lexicon and length also contribute to improve the basic classifiers. In contrast, rhetorical information alone modesty improves performance. This does not mean that RST is not a good indicator of subjectivity. In fact, some length features take advantage of RST information (i.e., length of the nuclei/satellite). Moreover, we conducted a preliminary analysis and found that some relations are highly indicative of subjectivity. For example, *attribution* recurrently appears in subjective sentences. In 30% of the subjective sentences this relationship occurs, whereas only 15% of the objective sentences contain an attribution relationship. Another important factor to take into account is that the presence of some RST relations is marginal. For instance, there are only 12 sentences with the *comparison* relation in our collection. This makes that some RST features are insignificant to discriminate between subjective and objective sentences. However, the inclusion of RST features seems to work well in combination with other evidence (e.g., combined with opinion lexicon features). This indicates that RST can modulate the influence of lexicon-based information. For instance, the presence of a *contrast* satellite could be valuable to increase the subjectivity score of a sentence. Consider the subjective sentence *"A degree of selfishness in capitalist countries seems to be part of the ideology, but one of the great lessons of this bloody 20th century was that pure self-interest needs to be tempered by a contribution to the more general good"*. This sentence has a *contrast* relationship in which the author compares the statements presented in the satellite (*"A degree ... ideology"*) and its nucleus (*"but one ... good"*). This is important evidence in favour of subjectivity (irrespective of the number of opinion terms in the sentence). On the other hand, the presence of a temporal satellite could be evidence in favour of objectivity. For example, the sentence *"Pakistan detonated a series of nuclear devices last month after India surprised the world with its tests"* has a *temporal* relationship between the nucleus (*"Pakistan ... month"*) and the satellite (*"after India ... tests"*). The *temporal* information provides the time when Pakistan detonated the nuclear devices, but the sentence does not express any opinion about it. Observe that with sentiment lexicon features alone, this sentence has a high probability of being classified as subjective because of the presence of opinionated words such as *surprised* or *detonated*. Note also that lexicon-based features lead to high-precision classifiers but the levels of recall are inferior to those achieved by other combination of features.

4.1 Feature Weights

After obtaining a linear SVM model, the weights (w_i) of the separating hyperplane can be used to assess the relevance of each feature [13]. The larger $|w_i|$ is, the more important the i_{th} is in the decision function of the SVM. Only linear SVM models have this indication, which naturally facilitates the analysis of the classifiers. This useful property has been used to gain knowledge of data and, for instance, to do feature selection [13,14]. A proper and direct comparison of the weights can only be done if all features are scaled into the same range. We therefore focus our analysis of feature weights on the *unigrams & bigrams + All* classifier after scaling the features into [0,1]. Table 4 presents the top 50 features ranked by decreasing absolute weight ($|w_i|$). A positive weight ($w_i > 0$) means that high values of the feature are indicative of the membership of the sentence into the subjective class. On the other hand, a negative weight ($w_i < 0$) means that high values of the feature are indicative of the membership of the sentence into the objective class. The two most discriminative features are the number of negative words and the position of the sentence in the document. Remember that the sentence position feature represents the order of a concrete sentence in its document (e.g., the third sentence of a document has a score of 3). A high w_i score makes that the final sentences have more chance of being labelled as subjective when compared to the initial sentences. This seems to indicate that writers tend to summarise their overall viewpoints at the end of the document. The most discriminative vocabulary features are the unigrams *objections* and *expressed*. The presence of these words in a sentence is a natural indicator of the opinionated nature of the sentence. Another interesting finding is that some of the most discriminative vocabulary features of the subjective class (i.e., unigrams or bigrams) are personal pronouns (e.g., they, I). These pronouns often appear in opinionated sentences to identify the holder of the opinion (e.g., *They are saying that the new product of company X is quite expensive*). On the other hand, the number of exclamations and interrogations has a high negative w_i score. This means that having interrogations or exclamations is indicative of objectivity in this dataset. This is intringuing, because the use of exclamations or interrogations have been associated to subjective content in the literature [4]. This outcome might be related with the nature of this repository (news articles). The use of interrogation or exclamations by journalists could be related to the way of writing to attract people's attention (e.g., open questions). This, however, requires further study and confirmation with other datasets.

If we focus our attention on the ranking of non-vocabulary features (i.e., all features but unigrams or bigrams), reported in Table 5, we observe that there are several RST features that help to detect subjective sentences. The most discriminative features tend to be terms provided by OF lexicon (i.e., Opinion feature set). This exemplifies the discriminative power of an opinion-based lexicon for classification purposes. Still, there are other interesting sets of features that help to understand the overall nature of the sentences. For instance, as we explained before, *evaluation*, *attribution* and *comparison* relationships are highly indicative of opinionated sentences (positive weights). For instance, it is common to find *attribution* statements when the author of the article writes about opinions of other people (e.g., *"According to the new CEO, the future of the company is brilliant"*). On the other hand *temporal* and *background* relationships seem to be indicative of objective sentences (negative weights). *background* statements help to comprehend the matter present in the nucleus. For instance, in the sentence, *"Culturally they are divided into peranakan and totok"*, the *background* satellite (*"Culturally"*) indicates the nature of the information presented in nucleus (*"they are divided into peranakan and totok"*). Additionally, *temporal* statements tend to be objective and are often used to locate events in time. For instance, in the sentence *"The day after the attacks, we saw immediate cancellations"*, the *temporal* satellite (*"The day after the attacks"*) indicates the precise period of time in which the action of the nucleus (*"we saw immediate cancellations"*) is performed.

Table 4. List of the 50 features with the highest $|w_i|$ in the best classifier(scaled). The features are ranked by decreasing $|w_i|$.

rank	w_i	feature	feature set	rank	w_i	feature	feature set
1	3.0439	#Neg	Opinion	26	-1.6449	to use	vocab.
2	2.4448	nSent	Position	27	1.6324	said in	vocab.
3	-2.4210	#ExcInt	Opinion	28	1.6182	programs	vocab.
4	2.3093	objections	vocab.	29	1.6095	ministers	vocab.
5	2.2380	expressed	vocab.	30	1.6087	US economy	vocab.
6	2.2355	they are	vocab.	31	1.5764	#Pos	Opinion
7	-2.2031	nSentsDoc	Length	32	1.5679	oil	vocab.
8	2.1838	globalisation	vocab.	33	1.5389	poor	vocab.
9	2.1239	actions	vocab.	34	1.5289	observers	vocab.
10	2.0839	Nor	vocab.	35	-1.5199	When	vocab.
11	2.0037	notably	vocab.	36	-1.5158	closer	vocab.
12	-1.9996	weather	vocab.	37	1.5126	terrorism	vocab.
13	1.9034	means	vocab.	38	1.5081	to have	vocab.
14	1.8829	something	vocab.	39	1.4957	leadership	vocab.
15	1.8137	I	vocab.	40	1.4938	looks	vocab.
16	-1.8026	market	vocab.	41	1.4859	#NegNorm	Opinion
17	-1.7575	expected	vocab.	42	-1.4843	see	vocab.
18	-1.7527	key	vocab.	43	1.4766	The economic	vocab.
19	-1.7205	will have	vocab.	44	-1.4742	with a	vocab.
20	1.7190	America	vocab.	45	1.4722	sufficient	vocab.
21	1.7002	#PosNorm	Opinion	46	1.4571	has a	vocab.
22	1.6894	should	vocab.	47	1.4555	mother	vocab.
23	1.6823	investors	vocab.	48	-1.4499	may be	vocab.
24	-1.6593	financial	vocab.	49	-1.4496	external	vocab.
25	-1.6522	world economy	vocab.	50	-1.4472	said Mr	vocab.

The results presented here are promising but the overall performance is still modest[7]. The nature of this collection (news articles), in which the authors (journalists) use an informative way of writing is challenging for sentiment detection algorithms. In other scenarios (e.g., blogs or opinion websites), the opinions are more explicit and this facilitates subjectivity classification. However, our results show that positional and discourse features are promising for developing new opinion classifiers able to overcome the limitations of classical vocabulary-based techniques. In fact, our techniques outperform content-based methods and popular opinion classifiers such as OF.

5 Related Work

Searching for relevant opinions within documents is a difficult task [15,8]. In [16] the authors considered the use of the first and the last sentences of a film review

[7] A random classifier for our imbalanced problem would get a $F1$ score around 33%.

Table 5. List of the non-vocabulary features with the highest $|w_i|$ in the best(scaled) classifier. The features are ranked by decreasing w_i.

rank	w_i	feature	feature set
1	3.0439	#Neg	Opinion
2	2.4448	nSent	Position
3	-2.4210	#ExcInt	Opinion
4	-2.2031	nSentsDoc	Length
5	1.7002	#PosNorm	Opinion
6	1.5764	#Pos	Opinion
7	1.4859	#NegNorm	Opinion
8	-1.4224	#ExcIntNorm	Opinion
9	1.3025	has *Evaluation* satellite	RST
10	-1.2566	nSentNorm	Position
11	0.9867	has *Attribution* satellite	RST
12	-0.8718	has *Temporal* satellite	RST
13	-0.8442	has *Background* satellite	RST
14	0.4591	has *Comparison* satellite	RST
15	0.4220	lengthSat	Length
16	-0.3927	has *Manner* satellite	RST
17	-0.3338	has *Cause* satellite	RST
18	-0.3034	lengthNuc	Length
19	-0.2612	has *Contrast* satellite	RST
20	0.2319	has *Condition* satellite	RST
21	-0.1997	has *Enablement* satellite	RST
22	0.1643	lengthSent	Length
23	-0.1635	has *Explanation* satellite	RST
24	-0.1170	has *Elaboration* satellite	RST
25	0.1112	has *Joint* satellite	RST
26	-0.0924	hasSat	RST

and evaluated their effect on accuracy. The impact of term positions in polarity classifiers was also studied in [17]. The results did not substantially differ with those obtained with no positional information. In this paper we have demonstrated that the use of positional information can lead to a better estimation of the subjectivity of the sentences in a news dataset.

Zirn et. al. [18] presented an automatic framework for fine-grained sentiment analysis at sub-sentence level in a product review scenario. Concretely, they used Markov logic to integrate polarity scores from different sentiment lexicons with information about relations between neighbouring segments of texts. They demonstrated that the use of rhetorical features improves the accuracy of polarity predictions. Somasundaran et al. [19] demonstrated the importance of general discourse analysis in polarity classification of multi-party meetings. The importance of RST for the classification of ambiguous sentences (i.e., sentences with conflicting opinions) was studied in [20]. In [21], the authors worked with film reviews and used RST to determine the importance of terms for polarity classification. With a sentence-level RST-analysis, they were able to outperform a document-level approach based on sentiment lexicons. However, they did not investigate the combination of RST and positional information and their solution works for a coarse-grained problem (document-level polarity estimation). In our paper we have demonstrated that RST can be applied at sentence level in combination with positional and other content-based features and this helps to select key subjective extracts from formal texts (news articles).

6 Conclusions

In this paper we explored the importance of sentence features such as positional or rhetorical features in fine-grained subjectivity classification processes. We demonstrated that these features are valuable and can be combined with more classical methods based on unigrams, bigrams and subjectivity lexicons. In the near future we would like to validate these findings against other datasets and study more advanced ways to combine features and classifiers.

Acknowledgments. This work was funded by *Secretaría de Estado de Investigación, Desarrollo e Innovación* from the Spanish Government under project TIN2012-33867.

References

1. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval **2**(1-2) (2007) 1–135
2. Liu, B.: Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers (2012)
3. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing contextual polarity in phrase-level sentiment analysis. In: Proc.of Human Language Technologies Conference/Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP 2005). (2005)
4. Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R.: Sentiment analysis of twitter data. In: Proceedings of the Workshop on Languages in Social Media. LSM '11, Stroudsburg, PA, USA, Association for Computational Linguistics (2011) 30–38
5. Mann, W.C., Thompson, S.A.: Rhetorical structure theory: Toward a functional theory of text organization. Text **8**(3) (1988) 243–281
6. Carlson, L., Marcu, D., Okurowski, M.E.: Building a discourse-tagged corpus in the framework of rhetorical structure theory. In: Proceedings of the Second SIGdial Workshop on Discourse and Dialogue - Volume 16. SIGDIAL '01, Stroudsburg, PA, USA, Association for Computational Linguistics (2001) 1–10
7. Seki, Y., Evans, D.K., Ku, L.W., Sun, L., Chen, H.H., Kando, N.: Overview of multilingual opinion analysis task at NTCIR-7. In: Proceedings of NTCIR-7. (2008)
8. Santos, R.L.T., He, B., Macdonald, C., Ounis, I.: Integrating proximity to subjective sentences for blog opinion retrieval. In: Proc. 31th European Conference on IR Research on Advances in Information Retrieval. ECIR '09, Berlin, Heidelberg, Springer-Verlag (2009) 325–336

9. Wiebe, J., Riloff, E.: Creating subjective and objective sentence classifiers from unannotated texts. In: Proceedings of the 6th international conference on Computational Linguistics and Intelligent Text Processing. CICLing'05, Berlin, Heidelberg, Springer-Verlag (2005) 486–497
10. Fan, R.E., Chang, K.W., Hsieh, C.J., Wang, X.R., Lin, C.J.: Liblinear: A library for large linear classification. J. Mach. Learn. Res. **9** (June 2008) 1871–1874
11. Nallapati, R.: Discriminative models for information retrieval. In: Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval. SIGIR '04, New York, NY, USA, ACM (2004) 64–71
12. Chenlo, J.M., Losada, D.E.: Effective and efficient polarity estimation in blogs based on sentence-level evidence. In: Proceedings of the 20th ACM international conference on Information and knowledge management. CIKM '11, New York, NY, USA, ACM (2011) 365–374
13. Chang, Y.W., Lin, C.J.: Feature ranking using linear svm. Journal of Machine Learning Research - Proceedings Track **3** (2008) 53–64
14. Brank, J., Grobelnik, M., Mili-frayling, N., Mladeni, D.: Feature selection using support vector machines. In: In Proc. of the 3rd Int. Conf. on Data Mining Methods and Databases for Engineering, Finance, and Other Fields. (2002) 84–89
15. Gerani, S., Carman, M.J., Crestani, F.: Proximity-based opinion retrieval. In: Proc. 33rd international ACM SIGIR conference on Research and development in information retrieval. SIGIR '10, New York, NY, USA, ACM (2010) 403–410
16. Pang, B., Lee, L.: A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Pr. of the ACL. (2004) 271–278
17. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? sentiment classification using machine learning techniques. In: Pr. of the Conference on Empirical Methods in Natural Language Processing. (2002)
18. Zirn, C., Niepert, M., Stuckenschmidt, H., Strube, M.: Fine-grained sentiment analysis with structural features. Number 12, Asian Federation of Natural Language Processing (2011)
19. Somasundaran, S., Namata, G., Wiebe, J., Getoor, L.: Supervised and unsupervised methods in employing discourse relations for improving opinion polarity classification. In: Proc. 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1 - Volume 1. EMNLP '09, Stroudsburg, PA, USA, ACL (2009) 170–179
20. Zhou, L., Li, B., Gao, W., Wei, Z., Wong, K.F.: Unsupervised discovery of discourse relations for eliminating intra-sentence polarity ambiguities. In: Proc. Conference on Empirical Methods in Natural Language Processing. EMNLP '11, Stroudsburg, PA, USA, ACL (2011) 162–171
21. Heerschop, B., Goossen, F., Hogenboom, A., Frasincar, F., Kaymak, U., de Jong, F.: Polarity analysis of texts using discourse structure. In: Proc. 20th ACM International Conference on Information and Knowledge Management. CIKM'11. CIKM '11, ACM press (2011) 1061–1070