Morgane Marchand, Romaric Besançon, Olivier Mesnard, Anne Vilnat

Domain Adaptation for Opinion Mining: A Study of Multi-polarity Words

Abstract

Expression of opinion depends on the domain. For instance, some words, called here multi-polarity words, have different polarities across domain. Therefore, a classifier trained on one domain and tested on another one will not perform well without adaptation. This article presents a study of the influence of these multi-polarity words on domain adaptation for automatic opinion classification. We also suggest an exploratory method for detecting them without using any label in the target domain. We show as well how these multi-polarity words can improve opinion classification in an open-domain corpus.

1 Introduction

With the advent of the Social Web, the way people express their opinions has changed: they can now post product reviews on merchant sites and express their point of views on almost anything in Internet forums, discussion groups, and blogs. Such online behaviour represents new and valuable sources of information with many practical applications. That is the reason why, in recent years, important research works have been undertaken on the subject of opinion mining. However, most works focus on how to characterize the opinion of texts in a given corpus, which is often domain-specific (*i.e.* the opinions in the texts are associated with the same type of objects), and little work have been done on words with different polarity across domains. Some words can indeed change their polarity between two different domains (Navigli, 2012; Yoshida et al., 2011). For example, the word "return" has a positive connotation in the sentence "I can't wait to return to my book". However, it can be seen as very negative when talking about some electronics device, as in "I had to return my phone to the store". This phenomenon happens even in more closely related domains: "I was laughing all the time" is a good point for a comedy but a bad one for a horror film. We call such words or expressions "multi-polarity words". This phenomenon is different from polysemy, as a word can keep the same meaning across domains while changing its polarity which can lead to classification error (Wilson et al., 2009). After a quick overview of the state of the art in this field, we present our study on these multi-polarity words. In section 3, we show that a significant amount of multi-polarity words influences the results of common automatic opinion classifiers. Their deletion or their differentiation leads to better classification results. We are also interested in the automatic detection of multi-polarity words when there is no annotation in the target domain. We propose a solution to solve this issue by using a set of common pivot words in order to compare distribution of candidate multi-polarity words in both domains. Finally, we show in section 4 that, even when a corpus does not contain explicit domain separation, the detection of multi-polarity words in implicit domains improves the opinion classification.

2 State of the art

Subjective expressions are words and phrases being used to express mental and emotional states like speculation, evaluation, sentiment or belief (Wiebe et al., 2005; Wiebe and Mihalcea, 2006; Wilson, 2008; Akkaya et al., 2009). They are called private states, that is to say, internal state which cannot be directly observed by others (Quirk and Crystal, 1985). On the contrary, polarity refers to positive or negative associations of a word or sense. Whereas there is a dependency in that most subjective senses have a relatively clear polarity, polarity can be attached to objective words or senses as well. Su and Markert (2008) give the example of the word tuberculosis: it does not describe a private state, is objectively verifiable and would not cause a sentence containing it to carry an opinion, but it does carry negative associations for the vast majority of people. Like Su and Markert (2008), we do not see polarity as a category that is dependent on prior subjectivity assignment and therefore applicable to subjective sense only. There is of course some correlations. A subjective sense of a word is likely to appear in a polar expression but can also appear in a neutral one. Similarly, an objective word can be used in a polar way.

Since a few years, interest on determining the polarity of ambiguous words has grown quickly (Wu and Jin, 2010). Practically all the existing annotation schemes for polarity include a "both" or "varied" flag (Su and Markert, 2008; Wilson et al., 2005). In their classification of the causes of variation in contextual polarity, Wilson et al. (2005) cite topic and domain. Moreover, in their study, Su and Markert (2008) notice that some preferences can exist depending on the domain or the topic of the text. They report 32.5 % of subjectivity ambiguous words in their corpus and the word sense disambiguation is not sufficient to remove the whole ambiguity. In Takamura et al. (2006, 2007), the authors propose latent variable model and lexical network to determine sentiment orientation of noun+adjective pairs. If the adjective is ambiguous, the classification is more difficult. Thus, the influence of domain on polarity is a very important field of research. In this study, we are looking for words or expressions (subjective or objective as well) which carry polar associations in a specific domain. Many of the words we are looking at would have no inherent polarity but can occur in polar contexts. We aim at imposing world knowledge and frequent discourse associations on these words.
This work is related to contextual or target polarity (Wilson et al., 2005; Fahrni and Klenner, 2008). Fahrni and Klenner (2008) focus on the target-specific polarity determination of adjectives. A domain specific noun is often modified by a qualifying

adjective. The authors argue that rather than having a prior polarity, adjectives are often bearing a target specific polarity. In some case, a single adjective even switches polarity depending on the accompanying noun. The authors use Wikipedia for automatic target detection and a bootstrapping approach to determine the target specific polarity of adjectives. They achieve good results but focus only on adjectives. On the contrary, Wilson et al. (2005) don't restrict them on adjectives but work only with phrases containing pre-determined clues. They focus on phrase-level sentiment analysis and first determine whether an expression is neutral or polar before disambiguating the polarity of the polar expression by using several rules and structural features.

In this study, we are interested in the influence of polarity-ambiguous words on polarity at text level. In state of the art, most works deal with a pre-existent lexicon of prior polarity. They aim at improving it, for example by weighting the different polarity of a word depending on the domain (Choi and Cardie, 2009). These particularized lexicons can then be used by a rule-based classifier (Ding et al., 2008).
As for studies on corpus-based only classifiers at text level, they focus mainly on the representation of data (Glorot et al., 2011; Huang and Yates, 2012). The adaptation error of a classifier depends indeed on its performance on the source domain and on the gap between source and target words distribution (Ben-David et al., 2007). With a good projection, a link can be established between the words of the target domain which are missing from the source domain and the other words (Pan et al., 2010; Blitzer et al., 2007). However, if a word in a text has different polarity in source and target domain, it will still introduce an error. So, identification of multi-polarity words is complementary to these approaches and their improvements can be combined. However, the influence of words with several polarities on automatic classifiers is rarely studied. One noticeable exception is the work of Yoshida et al. (2011). They use a bayesian formulation and focus more precisely on the influence of the number of source and target domains, using up to fourteen domains.

In all these works, the object of study can vary. For example, Wilson et al. (2009) use a pre-existing lexicon of polar words. The coverage of their lexicon is 75 % of the polar phrases of their corpus. On the contrary, Fahrni and Klenner (2008) focus on adjectives. In our study, we do not presume of what words or phrases are bearing polar information. We have chosen to automatically select them and classify them in one step. Therefore, we have to be attentive to avoid selecting peculiarities of the corpus.
As said before, we are working at text level. We are then interested on words or phrases which denote polarity at the text level. Some of them do not denote polarity at phrase level and then would not be considered by previous work. Among these words and phrases, we are interested only on those we call multi-polarity words. That is to say those which denote at text level a different polarity according to the general domain of the text.

3 A study of multi-polarity words

In this section, we present a study of multi-polarity words. The first part is dedicated to a qualitative and quantitative study of these words. In a second part, we present an estimation of their influence on an usual automatic classifier. Finally, we explore the detection of multi-polarity words without using any target label.

3.1 Description of the corpus

For this study, we have used the *Multi-Domain Sentiment Dataset*, collected by Blitzer et al. (2007). It contains four thematic corpora (*DVDs*, *kitchen*, *electronics* and *books*) of reviews collected on Amazon. Each corpus contains 1000 positive reviews, 1000 negative reviews and some unlabelled reviews. These reviews are represented with a bag of words of uni- and bi-grams. In this article, "word" is used to denote uni- or bi-grams.

3.2 Supervised detection of multi-polarity words

Multi-polarity words are first detected using a supervised approach, using the labelled reviews of each pair of thematic corpora. We make the common assumption that positive words will mostly appear in positive reviews and negative words in negative reviews. Then, for each word, we determine if its distribution in positive and negative reviews of target domain is statistically different or not from its distribution in positive and negative reviews of source domain[1]. For that purpose, we use a χ^2 test with a risk of false positive of 1%. The words are also selected only if they occur more often than a given threshold (minOcc) and if their difference of positivity between the two domains is higher than a second threshold (minDiff). These parameters are linked. If one of them is increased (less restrictive), the other one should be decreased (more restrictive) in order to keep the same level of performance. In a rank study, we have shown that they are approximatively linearly dependent.

Word	*region*	*I loved*	*worry*	*compare*	*return*
electronics	0.154	0.091	0.929	0.846	0.055
books	0.818	0.735	0.3	0.263	0.633

Table 1: Some example of percentage of presence in positive reviews for two domains. This score range from 0 (very negative) to 1 (very positive). A gap of 0.5 is then very significant (a neutral word becomes highly valued).

We present in Table 1 some multi-polarity words detected with this χ^2 test. As we detect our multi-polarity words based on a specific corpus, we have to be careful to

[1] Some words can have different polarity inside one domain but we only consider here the global polarity.

avoid selecting peculiarities of the corpus[2]. A more detailed analysis of this phenomenon leads us to the conclusion that words can change their polarity for multiple reasons. We propose the following classification of multi-polarity-words:

Corpus bias The change of polarity can be linked to a corpus bias: for instance, the word "superman" is very positive in the *books* corpus and negative in the *DVDs* corpus only because the film is often considered as a poor adaptation of the beloved comics.

Multiple word sense The multi-polarity of a word can be linked to polysemy. In "*I had to return my phone to the store*" or "*I can't wait to return to my book*", the word "return" has different polarities but also different senses. In this case, a pre-processing using word sense disambiguation methods or subjectivity word sense disambiguation methods like in (Akkaya et al., 2009) can be useful.

Relative quality Some adjectives or qualifiers without prior polarity can be positive or negative depending on their targeted object (Fahrni and Klenner, 2008). To be "*unpredictable*" is good for a film scenario but bad for a software.

Author's politic orientation Some words can change polarity depending the opinion of the writer. It often concerns political terms (e.g. "capitalism").

Comparison Comparative opinions ("*better than...*") are difficult to handle because the opinion characterization relies on the detection of which part of the comparison is the main subject. Some work has been developed about this specific problem (Ganapathibhotla and Liu, 2008). However, we have detected general habits in the different corpora. In *electronics* or *kitchen* corpora, comparisons are very common and in a huge majority, the topic of the review is in the first place of the comparison, whereas the opposite trend is found in *DVDs* or *books* corpora.

Temporal aspect The polarity of some words can be connected to an associated temporal information. For example, "*I loved this book*" is positive, however "*I loved this camera*" is usually negative because the camera doesn't work any more. "*I loved*" is therefore negative in *electronics* corpus, however, the present form "*I love*" stays positive.

Some of these categories can be handled other way, as *multiple word sense* or *comparison* categories. However, the effects of *relative quality* or *temporal aspect* can't be suppressed with usual treatment. That is why a study of these multi-polarity words is necessary.

3.3 Influence of multi-polarity words on automatic classifiers

The second part of our study on multi-polarity words aims at assessing the influence of these words on opinion classification tools based on machine learning techniques.

[2] A bigger manual evaluation is in progress.

For this purpose, we used a boosting method, BoosTexter (Schapire and Singer, 2000), because this method makes it easier to check which words are important for the classification. Indeed, words are chosen as weak classifiers, and if a word is selected early, it is considered very useful for the classification task. For each pair of corpora, we checked when the multi-polarity words where selected by Boostexter[3] as weak classifiers.

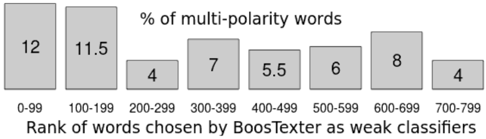

Figure 1: Average number of multi-polarity words among those selected by BoosTexter (with a step of 100 words), calculated on all the source-target pairs.

Figure 1 shows the average number of multi-polarity words for each 100 weak classifiers. Among the first 200 weak classifiers, 12% are multi-polarity words, which is relatively important and proves that a naïve handling of these words can generate noise in the opinion detection process. We highlighted this noisy influence with two simple experiments of features selection: the multi-polarity words are either deleted from the feature set or differentiated (by replacing a single *word* feature by two features *word-SOURCE*, *word-TARGET*).

	B-D	B-E	B-K	D-B	D-K	E-B	E-D	E-K	K-B
Normal	76.4	72.9	77.2	75.35	77.65	69.61	71.20	83.51	70.67
Diff.	**77.0**	**75.3**	**78.25**	75.35	*76.95*	**71.1**	**73.15**	*82.95*	**72.75**
Del.	76.25	**74.3**	**78.1**	**76.75**	*77.0*	**71.9**	**72.85**	*82.9*	**73.05**
	+0.6	+2.4	+1.05	+1.4	-0.7	+2.29	+2.05	-0.61	+2.38

	D-E	K-D	K-E
Normal	76.15	74.49	81.4
Diff.	75.7	74.3	81.35
Del.	75.8	74.65	82.1

Table 2: Accuracy for a BoosTexter classifier trained on a source domain and tested on a target domain (S-C); D : DVD, B : books, E : electronics, K : kitchen. Significant improvement are in bold and significant deterioration are in italic.

Table 2 shows that this very simple deletion (or differentiation) of multi-polarity words improves the classification for almost all the pairs Source-Target. Indeed, feature

[3]We have used the discrete AdaBoost.MH version, setting the number of iterations to 1000.

selection is known to be beneficial for domain adaptation (Satpal and Sarawagi, 2007). Moreover, a weighting of these multi-polarity words, rather than a complete deletion, is likely to give better results (Choi and Cardie, 2009). In a similar way, Akkaya et al. (2009) work on subjectivity ambiguous words and use subjectivity word sense disambiguation in order to improve contextual classification of polarity at sentence level. They remove subjective words used in objective context and the accurracy of their automatic classifier improves of three points.
These results justify the necessity of a dedicated handling of multi-polarity words, and of an automatic detection of these words in new domains.

3.4 Automatic detection of multi-polarity words in a new domain

The precedent study is based on a detection of multi-polarity words relying on annotations in both source and target domains. However, in a realistic application, the adaptation of a domain-specific opinion mining tool to a new domain has often to deal with no or few annotations of target domain. Automatic labelling can be useful but is not always possible. We present here our exploratory method for automatic detection of multi-polarity words without any target annotation.

3.4.1 Overview of the approach

The proposed method relies on a list of pivot words. They should belong to both source and target domain, be useful for the opinion classification task and have a stable polarity across domains. Their automatic selection is explained below. These pivot words are used in order to compare the distribution of the others words in source and target domains. For each word, and for each domain, we create its co-occurrence profile with respect to the list of pivot words. After that, a χ^2 test is applied to decide if, for a given word, its co-occurrence profiles in the source and target domains are statistically different (the word is considered as a multi-polarity word) or not (the word is then seen as a single-polarity word).

The pivot words are selected in two steps. First, a pre-selection is performed in order to keep only words which appear nearly as many time in both domains and are at the same time useful for opinion classification in source domain. Then, an iterative process removes from this list the words which have several polarities.
For the pre-selection step, we first compute, using only the annotated documents from the source domain, the mutual information $MI_{P,N}$ between the presence or absence of a word in a review and its positive/negative label. The selected pivot words should be useful for opinion classification and therefore have a high value for this mutual information score. We set a minimum threshold on this $MI_{P,N}$ in order to keep at least 1000 words. Following the same idea, we then compute, using the documents from both domains, the mutual information $MI_{S,T}$ between the presence and the absence of a word in a review and its source/target label. Words which are not specific to a domain should then have a low value for this mutual information score. The pivot

words candidates are ranked using this $MI_{S,T}$ and only the 1000 words with the lower values of $MI_{S,T}$ are kept.
After this pre-selection of pivot words, we detect the multi-polarity words among them using the same procedure as described in previous section but on pivot words themselves. We remove from the list the word which is the most likely to change polarity. Then, we iterate the process until no more pivot words are detected as multi-polarity words.

3.4.2 Evaluation of the results

This automatic method selects too many words. Therefore, in an in-context evaluation like in section 3.3, the accuracy drops drastically. In order to have an idea of the pertinence of our method, we have compared the words obtained automatically with our method (using only source labels) with those obtained by using labels of both source and target domains, as described in section 3.2. The automatic method selects more multi-polarity words (circa 1600 words) than the supervised one (circa 400 words), which explains the low precision score, as shown in table 3. Therefore, if all the detected words are deleted from the training corpus like in section 3.3, the accuracy is lower. However, precision can be increased without decreasing the recall by keeping only the words which are detected as multi-polarity words with the higher confidence: the values are presented in the column *max precision*. This confirms that our method indeed selects the multi-polarity words first: more work must be undertaken to find the optimal threshold for this selection.
Moreover, if we only consider multi-polarity words which are actually used by the classifiers (see figure 1), the average recall is 83.4 % for words selected in the first 100 weak classifiers (column *Recall 100*) and 71.3 % for the first 300 weak classifiers (column *Recall 300*). Therefore, the majority of multi-polarity words which are not detected are those with few influence on opinion classification.
So, despite a low precision, the results of our automatic detection method without using any target annotation are very promising.

Precision tot.	Recall tot.		Precision max.		Recall 100	Recall 300
16 %	60.5 %		18.1 %		83.4 %	71.3 %

Table 3: Comparison between words selected by the automatic method with those selected by the supervised one. The scores are the averaged recalls calculated on all the possible pairs Source-Target.

4 Use of multi-polarity words for open-domain opinion mining

In this section, we focus on another real case problem and present how to make use of the multi-polarity words in the context of opinion mining in open domain (i.e. in a general corpus that contains documents from different domains but without information on the domains). In this context, we cannot rely on the domain labels to detect multi-polarity words. We propose in this case to automatically find the different underlying domains of the documents in order to separate the general training corpus into smaller thematic corpora. Then, we apply the supervised detection method, presented in section 3.2, to detect multi-polarity words. These words are taken into account for learning several specific classifiers, one per thematic sub-corpus. The results of these classifiers are merged to produce the final opinion classification.

4.1 Overview of the method

To make use of the multi-polarity words in a labelled open domain corpus, we first have to extract the underlying domains in the documents and assign each document to a domain. We obtain several domains, not only two (one source and one target) like in the previous experiments. Therefore, we apply the supervised detection (3.2) of multi-polarity words several time, considering each domain versus all the others. We obtain as many multi-polarity words lists as underlying detected domains. For each multi-polarity words list, we create a new training corpus by deleting or differentiating the words of the list like in section 3.3. Opinion classifiers are created on these new training corpora. At last, we have one classifier per underlying detected domain. For classifying a new text, we merge the answers of the different classifiers according to the degree of relation of the new text to the underlying detected domains.

4.2 Evaluation

The evaluation of the proposed method is performed on the corpus of tweets from the task 2 of SemEval 2013 (Wilson et al., 2013). These tweets are separated in three classes: positive, negative and neutral. We use as training corpus the training data, merged with the development data and we balance the different classes. So, our final system is trained on 4500 tweets (1500 of each class, chosen randomly).
First, we remove the web addresses from the tweets to reduce the noise. Then, we extract the emoticons and use the number of occurrences of each type (smile, tears, heart...) as features. Finally, we perform a lemmatization of the text, using the linguistic analyser LIMA (Besançon et al., 2010). Table 4 shows a tweet example.

Bag of words features	Emoticon type feature
wow lady gaga be great	Smile 1

Table 4: "*WOW!!! Lady Gaga is great =)*"

4.2.1 Domain generation

As the corpus has no domain label, we first have to identify the underlying domains and assign a domain to each tweet. For that purpose, we use Latent Dirichlet Allocation (LDA) (Blei et al., 2003). LDA has already been used in aspect-based review analysis, which is close to our work. In (Titov and McDonald, 2008a,b), the authors introduce a model mixing global and local topics for aspect-based review analysis. They use the manual annotations of reviewers in order to improve the topics identification.
In our experiment on tweets, we chose the Mallet LDA implementation (McCallum, 2002).The framework uses Gibbs sampling to constitute the sample distributions which are exploited for the creation of the topic model. The model is built using the lemmatized tweets from the training and development data. We performed tests with different numbers of topics and the 5 topics version, presented in Table 5, appears to be the most efficient. Each tweet is then represented by a vector of length 5, where the i-th value is the proportion of words of the tweet which belong to the i-th topic.

Topic Film	tonight, watch, time, today
Topic Obama	win, vote, obama, black
Topic Sport	game, play, win, team
Topic Informatic	apple, international, sun, anderson
Topic Show	ticket, show, open, live

Table 5: Most representative words of each topic. We named the topics to make the presentation of data and results more readable.

Then, we subdivide the corpus in 5 sub-parts, or domain, each of them associated with one underlying detected topic. We have tested two types of subdivision. In the first one, called *all training tweets version*, a tweet is associated with its more related topic. For example, if its proportion of words belonging to the topic *sport* is 55 %, the tweet will be part of the sub-part associated to sport domain. In the second subdivision, called *domain confident training tweets version*, a tweet is taken into account only if more than 75 % of its words belong to the same topic. Therefore, the precedent example tweet will not be used. In this version, the sub-parts are more focused on only one topic. In return, they contain less training tweets (2889 tweets altogether).

4.2.2 Detection of multi-polarity words

For detecting the multi-polarity words, we use the positive and negative labels of the training data, as described in the section 3.2. We apply this detection for each sub-part. Each time, we detect the words which change their polarity between a specific sub-part of the training corpus and its complement (all the others tweets). For example, the word “black” is detected as positive in the second sub-part, related to the election of Barack Obama, and neutral in the rest of the tweets. At the end of this procedure, we have 5 collections of words which change their polarity (one different collection for each sub-part). These collections are rather small: from 21 to 61 multi-polarity words are detected according to the domain.

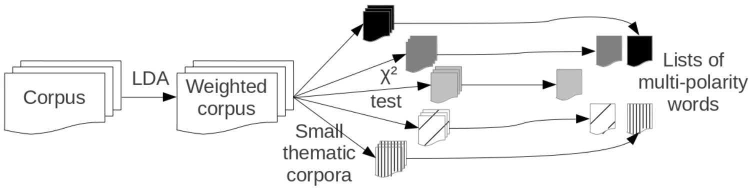

Figure 2: Detection of multi-polarity words after splitting the training corpus in 5 small thematic corpora using Latent Dirichlet Allocation.

4.2.3 Differentiation of multi-polarity words

After the automatic separation of the training corpus in different sub-parts associated to a specific domain and the detection of the words which change their polarity according to these domains, we integrate this knowledge in the opinion classifier. We produce a different corpus for each domain, by modifying the original one using the associated list of multi-polarity words. We then train a classifier on these modified corpus and obtain 5 domain-specific classifiers. As for the experiment described in section 3.3, we have tested two types of modification: differentiation or deletion. We have also performed a control experiment using only the separation into different domains but not the associated multi-polarity words. These modifications are described below:

Domain-specific version Different independent classifiers are trained on each domain-specific sub-part of the corpus, without any modification of the data. This is a control experiment. It uses only the domain information of the partitioning but not the multi-polarity words information.

Differentiation version Different domain-specific classifiers are trained on the whole corpus, modified like the experiments in section 3.2: each multi-polarity word for the domain X is replaced by a feature *word_X* in the sub-part of the corpus corresponding to this domain and left unchanged in the rest of the corpus. Hence, for each domain, we modify a different part of the original whole corpus.

Delete version Different domain-specific classifiers are trained. Each multi-polarity word for the domain X is removed from the whole corpus (different words are removed for the different domains, creating different versions of the corpus)

We then have 5 classifiers for classifying new tweets, each of them associated to one domain. Test tweets have no domain labels either. So, we first determine their topic

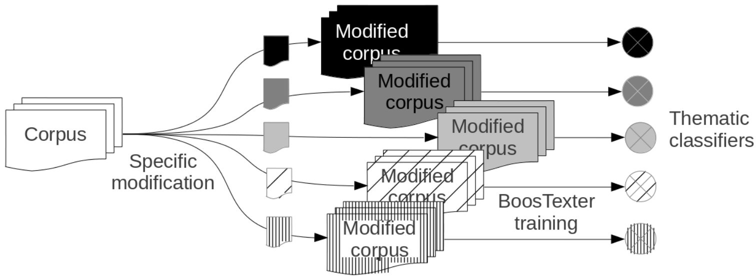

Figure 3: Flow of data. The modification is different for each version.

profile using LDA topic model. Then, we apply the 5 classifiers on the new tweet and obtain 5 answers. We use a mix of the 5 answers of the classifiers with weights according to the LDA mixture. This flow is presented in figure 3. We have tested several weighting schemes for this combination and the more efficient was the exponential of the LDA score.

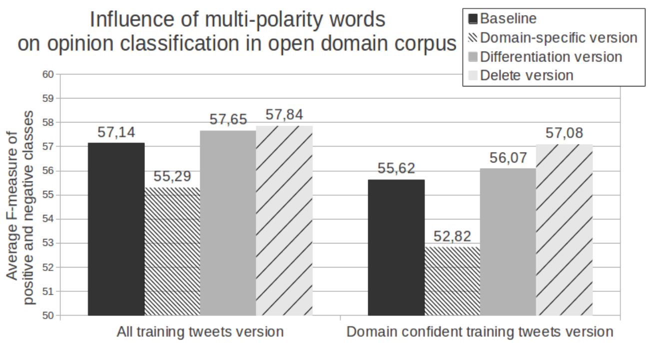

Figure 4: Average F-measure of positive and negative classes using two different training corpus: all training tweets or domain confident training tweets versions.

Figure 4 shows the results of these different integration of multi-polarity words using the two different initial training corpora created as described in section 4.2.1: all training tweets and domain confident training tweets versions.

4.3 Analysis of the results and discussion

We have described a method to include domain information in an open-domain corpus to improve opinion classification at text level. As we do not have reference domain label for the documents, we create a partition using a detection of the latent topics using LDA. The *Domain-specific* version, which does not take into account the multi-polarity words, degrades the performances(-1.85% in the first experiment, -2.8% in the second). We think it is due to the rather small size of some training sub-corpora of the partition. On the contrary, the results with all the versions which integrate multi-polarity words show an improvement of the F-measure. We have tested the significance of this improvement with a randomization test. In the case of the *Delete* version, the improvement is significant (p-value is 0.03). The final improvement is rather small, however, it has to be related to the small number of multi-polarity words we have detected (in average, 36 words per domain). We think that the considered collection of tweets chosen for the evaluation is too small for the χ_2 test to detect a lot of words with enough confidence. In comparison, in our experiment on reviews, we detected about 400 multi-polarity words per domain. It is also worth noticing that for the domain confident experiment, the improvement is more sensible (+1.46% versus +0.70%) even if the absolute value of the score is not better, due to a much smaller training data. Moreover, in this case, the significance of the *Delete* version is higher (with a p-value of 0.005). These results are very promising and show the interest of taking into account multi-polarity words.
Another issue for this method is its dependency on the approach which is chosen to separate the corpus into different domains. We used LDA for this purpose but we plan to test a more supervised method using Explicit Semantic Analysis (Gabrilovich and Markovitch, 2007) and based on the categories of Wikipedia, in order to have more control on the domains (i.e. propose general domains that are not corpus-dependent).

5 Conclusion

In this article, we have studied the influence of multi-polarity words on the performance of the automatic classification of opinion at text level. We have shown that these words are frequent and have influence on the performance of automatic classifiers in a corpus of domain-specific reviews and in an open-domain corpus of tweets. At the present time, a manual evaluation of these words is in progress. We discussed the real case where there is no labels available in the target domain and present an exploratory method for detecting multi-polarity words using carefully selected pivot words. Then we showed that the detection of multi-polarity words is also useful in an open-domain corpus. Further works will be made on the selection criteria of multi-polarity words, especially in the case where no target label is used.

References

Akkaya, C., Wiebe, J., and Mihalcea, R. (2009). Subjectivity word sense disambiguation. In *EMNLP*.

Ben-David, S., Blitzer, J., Crammer, K., and Pereira, F. (2007). Analysis of representations for domain adaptation. *Advances in neural information processing systems*, 19:137.

Besançon, R., el de Chalendar, G., Ferret, O., Gara, F., Mesnard, O., Laïb, M., and Semmar, N. (2010). Lima : A multilingual framework for linguistic analysis and linguistic resources development and evaluation. In *Proceedings of LREC'10*.

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

Blitzer, J., Dredze, M., and Pereira, F. (2007). Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In *ACL*.

Choi, Y. and Cardie, C. (2009). Adapting a polarity lexicon using integer linear programming for domain-specific sentiment classification. In *EMNLP*.

Ding, X., Liu, B., and Yu, P. S. (2008). A holistic lexicon-based approach to opinion mining. In *Proceedings of WSDM*, pages 231–240. ACM.

Fahrni, A. and Klenner, M. (2008). Old wine or warm beer: Target-specific sentiment analysis of adjectives. In *Symposion on Affective Language in Human and Machine, AISB Convention*.

Gabrilovich, E. and Markovitch, S. (2007). Computing semantic relatedness using wikipedia-based explicit semantic analysis. In *IJCAI*, volume 7, pages 1606–1611.

Ganapathibhotla, M. and Liu, B. (2008). Mining opinions in comparative sentences. In *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*, pages 241–248. Association for Computational Linguistics.

Glorot, X., Bordes, A., and Bengio, Y. (2011). Domain adaptation for large-scale sentiment classification: A deep learning approach. In *ICML*.

Huang, F. and Yates, A. (2012). Biased representation learning for domain adaptation. In *EMNLP*, pages 1313–1323, Jeju Island, Korea. Association for Computational Linguistics.

McCallum, A. K. (2002). Mallet: A machine learning for language toolkit.

Navigli, R. (2012). A quick tour of word sense disambiguation, induction and related approaches. *SOFSEM 2012: Theory and Practice of Computer Science*, pages 115–129.

Pan, S., Ni, X., Sun, J., Yang, Q., and Chen, Z. (2010). Cross-domain sentiment classification via spectral feature alignment. In *WWW*, pages 751–760. ACM.

Quirk, R. and Crystal, D. (1985). *A comprehensive grammar of the English language*, volume 6. Cambridge Univ Press.

Satpal, S. and Sarawagi, S. (2007). Domain adaptation of conditional probability models via feature subsetting. In *PKDD*.

Schapire, R. and Singer, Y. (2000). Boostexter: A boosting-based system for text categorization. *Machine learning*, 39(2):135–168.

Su, F. and Markert, K. (2008). From words to senses: a case study of subjectivity recognition. In *International Conference on Computational Linguistics*.

Takamura, H., Inui, T., and Okumura, M. (2006). Latent variable models for semantic orientations of phrases. In *EACL*.

Takamura, H., Inui, T., and Okumura, M. (2007). Extracting semantic orientations of phrases from dictionary. In *HLT-NAACL*, pages 292–299.

Titov, I. and McDonald, R. (2008a). A joint model of text and aspect ratings for sentiment summarization. In *ACL*.

Titov, I. and McDonald, R. (2008b). Modeling online reviews with multi-grain topic models. In *WWW*.

Wiebe, J. and Mihalcea, R. (2006). Word sense and subjectivity. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pages 1065–1072. Association for Computational Linguistics.

Wiebe, J., Wilson, T., and Cardie, C. (2005). Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, 39(2-3):165–210.

Wilson, T., Kozareva, Z., Nakov, P., Ritter, A., Rosenthal, S., and Stoyanov, V. (2013). Semeval-2013 task 2: Sentiment analysis in twitter. In *7th International Workshop on Semantic Evaluation*.

Wilson, T., Wiebe, J., and Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *HLT/EMNLP*.

Wilson, T., Wiebe, J., and Hoffmann, P. (2009). Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis. *Computational Linguistics*, 35:339–433.

Wilson, T. A. (2008). *Fine-grained subjectivity and sentiment analysis: recognizing the intensity, polarity, and attitudes of private states*. ProQuest.

Wu, Y. and Jin, P. (2010). Semeval-2010 task 18: Disambiguating sentiment ambiguous adjectives. In *5th International Workshop on Semantic Evaluation*, pages 81–85.

Yoshida, Y., Hirao, T., Iwata, T., Nagata, M., and Matsumoto, Y. (2011). Transfer learning for multiple-domain sentiment analysis - identifying domain dependent/independent word polaritys. In *AAAI*.