# UA-ZBSA: A Headline Emotion Classification through Web Information

**Zornitsa Kozareva, Borja Navarro, Sonia Vázquez, Andrés Montoyo**
DLSI, University of Alicante
Carretera de San Vicente S/N
Alicante, Spain
03080
zkozareva,borja,svazquez,montoyo@dlsi.ua.es

## Abstract

This paper presents a headline emotion classification approach based on frequency and co-occurrence information collected from the World Wide Web. The content words of a headline (nouns, verbs, adverbs and adjectives) are extracted in order to form different bag of word pairs with the joy, disgust, fear, anger, sadness and surprise emotions. For each pair, we compute the Mutual Information Score which is obtained from the web occurrences of an emotion and the content words. Our approach is based on the hypothesis that group of words which co-occur together across many documents with a given emotion are highly probable to express the same emotion.

## 1 Introduction

The subjective analysis of a text is becoming important for many Natural Language Processing (NLP) applications such as Question Answering, Information Extraction, Text Categorization among others (Shanahan et al., 2006). The resolution of this problem can lead to a complete, realistic and coherent analysis of the natural language, therefore major attention is drawn to the opinion, sentiment and emotion analysis, and to the identification of beliefs, thoughts, feelings and judgments (Quirk et al., 1985), (Wilson and Wiebe, 2005).

The aim of the Affective Text task is to classify a set of news headlines into six types of emotions: “anger”, “disgust”, “fear”, “joy”, “sadness” and “surprise”. In order to be able to conduct such multi-category analysis, we believe that first we need a comprehensive theory of what a human emotion is, and then we need to understand how the emotion is expressed and transmitted within the natural language. These aspects rise the need of syntactic, semantic, textual and pragmatic analysis of a text (Polanyi and Zaenen, 2006). However, some of the major drawbacks in this field are related to the manual or automatic acquisition of subjective expressions, as well as to the lack of resources in terms of coverage.

For this reason, our current emotion classification approach is based on frequency and co-occurrence bag of word counts collected from the World Wide Web. Our hypothesis is that words which tend to co-occur across many documents with a given emotion are highly probable to express this emotion.

The rest of the paper is organized as follows. In Section 2 we review some of the related work, in Section 3 we describe our web-based emotion classification approach for which we show a walk-through example in Section 4. A discussion of the obtained results can be found in Section 5 and finally we conclude in Section 6.

## 2 Related work

Our approach for emotion classification is based on the idea of (Hatzivassiloglou and McKeown, 1997) and is similar to those of (Turney, 2002) and (Turney and Littman, 2003). According to Hatzivassiloglou and McKeown (1997), adjectives with the same polarity tended to appear together. For example the negative adjectives “corrupt and brutal” co-occur very often.

The idea of tracing polarity through adjective co-occurrence is adopted by Turney (2002) for the binary (positive and negative) classification of text reviews. They take two adjectives, for instance "excellent" and "poor" in a way that the first adjective expresses positive meaning, meanwhile the second one expresses negative. Then, they extract all adjectives from the review text and combine them with "excellent" and "poor". The co-occurrences of these words are searched on the web, and then the Mutual Information score for the two groups of adjectives is measured. When the adjective of the review appear more often with "excellent", then the review is classified as positive, and when the adjectives appear more often with "poor", then the review is classified as negative.

Following Hatzivassiloglou and McKeown (1997) and Turney (2002), we decided to observe how often the words from the headline co-occur with each one of the six emotions. This study helped us deduce information according to which "birthday" appears more often with "joy", while "war" appears more often with "fear".

Some of the differences between our approach and those of Turney (2002) are mentioned below:

- objectives: Turney (2002) aims at binary text classification, while our objective is six class classification of one-liner headlines. Moreover, we have to provide a score between 0 and 100 indicating the presence of an emotion, and not simply to identify what the emotion in the text is. Apart from the difficulty introduced by the multi-category classification, we have to deal with a small number of content words while Turney works with large list of adjectives.

- word class: Turney (2002) measures polarity using only adjectives, however in our approach we consider the noun, the verb, the adverb and the adjective content words. The motivation of our study comes from (Polanyi and Zaenen, 2006), according to which each content word can express sentiment and emotion. In addition to this issue we saw that most of the headlines contain only nouns and verbs, because they express objectivity.

- search engines: Turney (2002) uses the Altavista web browser, while we consider and combine the frequency information acquired from three web search engines.

- word proximity: For the web searches, Turney (2002) uses the NEAR operator and considers only those documents that contain the adjectives within a specific proximity. In our approach, as far as the majority of the query words appear in the documents, the frequency count is considered.

- queries: The queries of Turney (2002) are made up of a pair of adjectives, and in our approach the query contains the content words of the headline and an emotion.

There are other emotion classification approaches that use the web as a source of information. For instance, (Taboada et al., 2006) extracted from the web co-occurrences of adverbs, adjectives, nouns and verbs. Gamon and Aue (2005) were looking for adjectives that did not co-occur at sentence level. (Baroni and Vegnaduzzo, 2004) and (Grefenstette et al., 2004) gathered subjective adjectives from the web calculating the Mutual Information score.

Other important works on sentiment analysis are those of (Wilson et al., 2005) and (Wiebe et al., 2005; Wilson and Wiebe, 2005), who used linguistic information such as syntax and negations to determine polarity. Kim and Hovy (2006) integrated verb information from FrameNet and incorporated it into semantic role labeling.

## 3 Web co-occurrences

In order to determine the emotions of a headline, we measure the Pointwise Mutual Information (MI) of $e_i$ and $cw_j$ as $MI(e_i, cw_j) = \log_2 \frac{hits(e_i,cw_j)}{hits(e_i)hits(cw_j)}$, where $e_i \in \{anger, disgust, fear, joy, sadness, surprise\}$ and $cw_j$ are the content words of the headline $j$. For each headline, we have six MI scores which indicate the presence of the emotion. MI is used in our experiments because it provides information about the independence of an emotion and a bag of words.

To collect the frequency and co-occurrence counts of the headline words, we need large and massive data repositories. To surmount the data sparsity problem, we used as corpus the World Wide Web which is constantly growing and daily updated.

Our statistical information is collected from three web search engines: MyWay[1], AlltheWeb[2] and Yahoo[3]. It is interesting to note that the emotion distribution provided by each one of the search engines for the same headline has different scores. For this reason, we decided to compute an intermediate MI score as $aMI = \frac{\sum_{s=1}^{n} MI(e_i, cw_j)}{s}$.

In the trail data, besides the MI score of an emotion and all headline content words, we have calculated the MI for an emotion and each one of the content words. This allowed us to determine the most sentiment oriented word in the headline and then we use this predominant emotion to weight the association sentiment score for the whole text. Unfortunately, we could not provide results for the test data set, due to the high number of emotion-content word pairs and the increment in processing time and returned responses of the search engines.

## 4 Example for Emotion Classification

As a walk through example, we use the *Mortar assault leaves at least 18 dead* headline which is taken from the trial data. The first step in our emotion classification approach consists in the determination of the part-of-speech tags for the one-liner. The non-content words are stripped away, and the rest of the words are taken for web queries. To calculate the MI score of a headline, we query the three search engines combining "mortar, assault, leave, dead" with the anger, joy, disgust, fear, sadness and surprise emotions. The obtained results are normalized in a range from 0 to 100 and are shown in Table 1.

| | **MyWay** | **AllWeb** | **Yahoo** | **Av.** | **G.Stand.** |
|---|---|---|---|---|---|
| **anger** | 19 | 22 | 24 | 22 | 22 |
| **disgust** | 5 | 6 | 7 | 6 | 2 |
| **fear** | 44 | 50 | 53 | 49 | 60 |
| **joy** | 15 | 19 | 20 | 18 | 0 |
| **sadness** | 28 | 36 | 36 | 33 | 64 |
| **surprise** | 4 | 5 | 6 | 5 | 0 |

Table 1: Performance of the web-based emotion classification for a trail data headline

[1] www.myway.com

[2] www.alltheweb.com

[3] www.yahoo.com

As can be seen from the table, the three search engines provide different sentiment distribution for the same headline, therefore in our final experiment we decided to calculate intermediate MI. Comparing our results to those of the gold standard, we can say that our approach detects significantly well the fear, sadness and angry emotions.

## 5 Results and Discussion

Table 2 shows the obtained results for the affective test data. The low performance of our approach is explainable by the minimal knowledge we have used. An interesting conclusion deduced from the trail and test emotion data is that the system detects better the negative feelings such as anger, disgust, fear and sadness, in comparison to the positive emotions such as joy and surprise. This makes us believe that according to the web most of the word-emotion combinations we queried are related to the expression of negative emotions.

| **UA-ZBSA** | **Fine-grained** | **Coarse-grained** | | |
|---|---|---|---|---|
| | **Pearson** | **Acc.** | **P.** | **R.** |
| **Anger** | 23.20 | 86.40 | 12.74 | 21.66 |
| **Disgust** | 16.21 | 97.30 | 0.00 | 0.00 |
| **Fear** | 23.15 | 75.30 | 16.23 | 26.27 |
| **Joy** | 2.35 | 81.80 | 40.00 | 2.22 |
| **Sadness** | 12.28 | 88.90 | 25.00 | 0.91 |
| **Surprise** | 7.75 | 84.60 | 13.70 | 16.56 |

Table 2: Performance of the web-based emotion classification for the whole test data set

In the test run, we could not apply the emotion-word weighting, however we believe that it has a significant impact over the final performance. Presently, we were looking for the distribution of all content words and the emotions, but in the future we would like to transform all words into adjectives and then conduct web queries.

Furthermore, we would like to combine the results from the web emotion classification with the polarity information given by SentiWordNet[4]. A-priory we want to disambiguate the headline content words and to determine the polarities of the words and their corresponding senses. For instance, the adjective "new" has eleven senses, where new#a#3 and new#a#5 express negativism, new#a#4 and new#a#9 positivism and the rest of the senses are objective.

[4] http://sentiwordnet.isti.cnr.it/

So far we did not consider the impact of valence shifter (Polanyi and Zaenen, 2006) and we were unable to detect that a negative adverb or adjective transforms the emotion from positive into negative and vice versa. We are also interested in studying how to conduct queries not as a bag of words but bind by syntactic relations (Wilson et al., 2005).

## 6 Conclusion

Emotion classification is a challenging and difficult task in Natural Language Processing. For our first attempt to detect the amount of angry, fear, sadness, surprise, disgust and joy emotions, we have presented a simple web co-occurrence approach. We have combined the frequency count information of three search engines and we have measured the Mutual Information score between a bag of content words and emotion.

According to the yielded results, the presented approach can determine whether one sentiment is predominant or not, and most of the correct sentiment assignments correspond to the negative emotions. However, we need to improve the approach in many aspects and to incorporate more knowledge-rich resources, as well as to tune the 0-100 emotion scale.

## Acknowledgements

This research has been funded by QALLME number FP6 IST-033860 and TEX-MESS number TIN2006-15265-C06-01.

## References

Marco Baroni and Stefano Vegnaduzzo. 2004. Identifying subjective adjectives through web-based mutual information. In Ernst Buchberger, editor, *Proceedings of KONVENS 2004*, pages 17–24.

Michael Gamon and Anthony Aue. 2005. Automatic identification of sentiment vocabulary: exploiting low association with known sentiment terms. In *Proceedings of the Workshop on Feature Engineering for Machine Learning in Natural Language Processing (ACL 2005)*, pages 57–64.

Gregory Grefenstette, Yan Qu, James G. Shanahana, and David A. Evans. 2004. Coupling niche browsers and affect analysis for an opinion mining application. In *Proceeding of RIAO-04*.

Vasileios Hatzivassiloglou and Kathleen R. McKeown. 1997. Predicting the semantic orientation of adjectives. In *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics (EACL)*.

Soo-Min Kim and Eduard Hovy. 2006. Extracting opinions, opinion holders, and topics expressed in online news media text. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text*, pages 1–8.

Livia Polanyi and Annie Zaenen. 2006. Contextual valence shifter. In James G. Shanahan, Yan Qu, and Janyce Wiebe, editors, *Computing Attitude and Affect in Text: Theory and Applications*, chapter 1, pages 1–10. Springer.

R. Quirk, S. Greenbaum, G. Leech, and J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman.

James G. Shanahan, Yan Qu, and Janyce Wiebe. 2006. *Computing Attitude and Affect in Text: Theory and Applications*. Springer.

Maite Taboada, Caroline Anthony, and Kimberly Voll. 2006. Methods for creating semantic orientation databases. In *Proceeding of LREC-06, the 5th International Conference on Language Resources and Evaluation*, pages 427–432.

Peter D. Turney and Michael L. Littman. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems*, 21(4):315–346.

Peter D. Turney. 2002. Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 417–424.

Janyce Wiebe, Theresa Wilson, and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation (formerly Computers and the Humanities)*, 39(2-3):165–210.

Theresa Wilson and Janyce Wiebe. 2005. Annotating attributions and private states. In Ann Arbor, editor, *Proceedings of the Workshop on Frontiers in Corpus Annotation II: Pie in the Sky*, pages 53–60.

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 347–354.