A CCG-based Approach to Fine-Grained Sentiment Analysis in Microtext

Phillip Smith & Mark Lee
School of Computer Science
University of Birmingham
Edgbaston, Birmingham, B15 2TT
{P.Smith.7, M.G.Lee}@cs.bham.ac.uk

Abstract

In this paper, we present a Combinatory Categorial Grammar (CCG) based approach to the classification of emotion in microtext. We develop a method that makes use of the notion put forward by Ortony, Clore, and Collins (1988), that emotions are valenced reactions. This hypothesis sits central to our system, in which we adapt contextual valence shifters to infer the emotional content of a text. We integrate this with an augmented version of WordNet-Affect, which acts as our lexicon. Finally, we experiment with a corpus of headlines proposed in the 2007 SemEval Affective Task (Strapparava and Mihalcea 2007) as our microtext corpus, and by taking the other competing systems as a baseline, demonstrate that our approach to emotion categorisation performs favourably.

Introduction

Text, no matter the length, can potentially convey an emotional meaning. As the availability of digitized documents has increased over the past decade, so the ability and need to classify this data by its affective content has increased. This in turn has generated a large amount of interest in the field of Sentiment Analysis.

Typical approaches to Sentiment Analysis tend to focus on the binary classification problem of valence: whether a text has a positive or negative sentiment associated with it. The task of classifying text by its valence has been applied successfully across varying datasets, from product reviews (Blitzer, Dredze, and Pereira 2007) and online debates (Mukherjee and Liu 2012), even spanning as far as the sentiment communicated through patient discourse (Smith and Lee 2012). While numerous works concentrate on the binary-classification task, the next logical task in sentiment analysis, emotion classification, can sometimes be overlooked, for numerous reasons.

Emotion classification provides a more complex problem than the polarity based sentiment analysis task. While both suffer from the subtleties that the implicit nature of language holds, one of the central reasons for its complexity is that there are a greater number of categories, emotions, in which to undertake classification. Additionally, there is no fixed number of categories, as varying theories of emotion have been proposed, each detailing a slightly different subset of emotions.

This paper will provide a general approach to emotion classification, which utilises the lexical semantics of words and their combinations in order to classify a text. We will experiment with our proposed method on the SemEval 2007 Affective Task, proposed by Strapparava and Mihalcea (2007). The task offered an interesting challenge for sentiment analysis, as little data was given for training, so supervised machine learning approaches that are common to text classification on the whole, were discouraged. This therefore encouraged competing systems to consider the syntax and semantics of language when crafting their approaches to classification. The task was split into two tracks, one for traditional valence classification, and one for emotion classification. Our system experiments with the latter track.

The SemEval Data Sets and Evaluation

The corpus that was compiled for the Affective Task consisted of general news headlines obtained from websites such as Google News and CNN. Whilst a corpus of headlines is not typical for sentiment analysis, this domain was chosen for the task in hand due to the salience of the emotions that are conveyed through the use of only a few thought provoking words. It is usual for sentiment analysis to be carried out on large document sets, where documents may consist of numerous paragraphs, but in the case of this task, sentiment analysis focused on the sentence level.

The headlines provided in the corpus were annotated by six independent annotators. Six different emotions that correspond with those proposed by Ekman (1982) were used as the category labels. These six emotions were *anger*, *disgust*, *fear*, *joy*, *sadness* and *surprise*. For each emotional category, the headline was annotated on a fine-grained scale between 0 and 100, dependent upon how strongly an annotator felt that a particular emotion was expressed. For the coarse-grained evaluations of systems, each emotion was mapped to a 0/1 classification, where 0=[0,50] and 1=[50,100].

The dataset that was released consisted of two sections, a trial set and a test set. The trial set, consisted of 250 headlines, and the test set, used for evaluating the systems consisted of 1,000 annotated headlines.

Outline of Our Approach

A central part of our approach to emotion classification was the use of an appropriate lexicon. Whilst a number of lexica for sentiment analysis exist such as SentiWordNet (Esuli and Sebastiani 2006) and AFINN (Hansen et al. 2011), as is the case with most approaches to sentiment analysis, valence is focused on, and emotions unfortunately are not considered. Therefore, in our approach to emotion classification, we use the optional lexicon of emotion bearing unigrams, WordNet-Affect, provided by the task organisers. This lexicon presents a mapping from emotional terms to the relevant emotional categories that were used to annotate the headlines in the affective task.

The WordNet-Affect dictionary alone would not suffice in a classification task from a specific genre of texts, namely headlines. WordNet-Affect contains hypernymic words associated with basic emotional concepts, but does not contain some of the more general emotion causing lexical items that are associated with headlines, such as *war*. Due to this, expansion of the lexicon with further emotion-bearing concepts was required.

Alongside the expansion of the lexicon, another occurrence in sentences needed to be taken into account: contextual valence shifters. For example, consider the sentence from the trial data set *'Budapest calm after night of violent protests'*. A basic bag-of-words approach to this may view the words (*violent, protests*) as fear, anger or sadness, whereas the only word that suggests joy is (*calm*). With a uniform scoring system in place, this headline would be incorrectly classified.

To overcome this short-coming in bag-of-words approaches to classification, sentence level valence shifters (Polanyi and Zaenen 2006) are implemented. These influential lexical items act by altering the valence of words around them. The combination of *calm after* suggests a change in valence of the sentence, and so the phrase *night of violent protests* is shifted from a negative to positive valence.

To apply this valence shifting technology to emotion classification, we must build upon the hypothesis proposed by Ortony, Clore, and Collins (1988) that emotions are rooted with either a positive or negative valence, and that most words have the capability to shift valence under certain contexts. In the case of this task, we assume only joy to be associated with a positive valence, and the emotions of anger, fear, disgust, sadness and surprise stem from a negative valence. In doing this, we are able to make fine-grained emotional classifications on the headlines.

In order to implement the contextual valence shifters, a relevant parser was required that could capture adequately the functionality of valence shifting lexical entities. The Categorial Combinatory Grammar (Steedman 2000) takes advantage of surface syntax as an interface to the underlying compositional semantics of a language, and therefore is suitable for discovering valence shifting terms. To intergrate the CCG formalism into our system, Clark and Curran's (Clark and Curran 2004) implementation of the parser was used.

Resources

To develop our system three main elements were integrated to tackle the problem of emotion classification:

- A lexicon of emotion bearing unigrams - an augmented version of WordNet-Affect
- Contextual Valence Shifters
- A Combinatory Categorial Grammar parser

These will further be described below.

WordNet-Affect

WordNet-Affect (Strapparava and Valitutti 2004) is a lexical resource developed by extending WordNet (Fellbaum 1998) with affective domain labels in order to produce a lexicon capable of associating affective concepts with affective words. To achieve this, WordNet-Affect (WN-A) introduces a hierarchy of affective labels whereby the included synsets are considered due to the affective concepts associated with them. This hierarchical emotional structure is modelled upon the hypernymic relations of WordNet. The affective domain labels (a-labels) consist of a number of concepts associated with affect, which include aspects such as emotion, mood, attitude and cognitive state. For the SemEval Affective Task, a subset of WN-A was released that specifically related to the six emotion categories that were used. An overview of this is given in the following table.

	Nouns	Verbs	Adjectives	Adverbs
Anger	99	64	119	35
Disgust	6	22	34	10
Fear	43	65	96	26
Joy	149	122	203	65
Sadness	64	25	169	43
Surprise	8	28	41	13

Table 1: WordNet-Affect word counts

Contextual Valence Shifters

A prevalent aspect of language is that the lexical choice of the writer is salient in conveying attitude. However, as Polanyi and Zaenen (2006) point out, the base valence of a lexical item is often modified by the polarity of its neighbouring terms, and this is something that is often overlooked in the sentiment analysis literature. For example, in the phrase *'she is not happy'*, the use of the word *not* shifts the valence of the term *happy* from a positive valence to a negative one.

However, whilst the valence may shift polarity, the same cannot be said for the emotion in the example phrase. An assumption is to uniformly shift an emotion to its presumed opposite emotion, in this case, sadness. There lies a problem with this though, as *'she is not happy'* is not equivalent to *'she is sad'*. A number of different emotions that are negatively valenced, such as anger, could be inferred from the original example sentence. Due to this, the use of the hypothesis put forward by Ortony, Clore, and Collins (1988)

$$\frac{\underset{\lambda x.female(x)}{\overline{\text{Female}}} \quad \underset{astronaut}{\overline{\text{astronaut}}}}{female(astronaut)} \quad \frac{\underset{\lambda x.\lambda y.sets(x)(y)}{\overline{\text{sets}}} \quad \frac{\underset{\lambda x.new(x)}{\overline{\text{new}}} \quad \underset{record}{\overline{\text{record}}}}{new(record)}}{\lambda y.set(new(record))(y)}$$

$$sets(new(record))(female(astronaut))$$

Figure 1: The semantic structure of the sentence: 'Female astronaut sets new record'

is key in determining an overall shift in emotion within a phrase or sentence.

Lexical items such as "very" and not" can be used under a variety of emotional settings, but their main role is to contribute to the strength of the resulting emotion or emotions that are conveyed within a sentence.

Combinatory Categorial Grammar

Combinatory Categorial Grammar (CCG) (Steedman 2000) is a popular grammar formalism that builds upon combinatory logic in order to undertake efficient natural language parsing. The formalism is based upon the notion that in natural language the surface syntax acts as an interface to the underlying compositional semantics of a language.

CCGs map lexical elements of a sentence, such as nouns and adjectives, to a syntactic category. In addition to these mappings, the CCG formalism also offers a variety of combinatory rules, such as coordination and type-raising, that specify how constituent categories can be combined into larger chunks in order to provide a suitable parse for a sentence, or fragment of a sentence.

The CCG formalism provides two types of syntactic category: primitive and complex. The primitive category is recursively defined as the set of terms that include basic categories such as *V* (verb), *VP* (verb phrase), *S* (sentence) and so on. Complex categories act as functions within the grammar, and are compounds of the primitive categories. They typically take the form A/B or $A\backslash B$, where A and B are primitive categories. In this notation, the argument appears to the right on the slash, and the resulting category appears to the left of the slash. So, in the previous example, B is the argument given to the function, and A is the resulting category. The directionality of the slash indicates which side of the functor the argument must appear on. A forward slash indicates that the argument must appear to the right of the given constituent, while a backslash indicates that it must appear to the left. The following example shows how constituents combine in order to give a full parse of the sentence 'Female astronaut sets new record ':

$$\frac{\underset{N/N}{\overline{\text{Female}}} \quad \underset{N}{\overline{\text{astronaut}}}}{\cfrac{N}{NP}} \quad \frac{\underset{(S\backslash NP)/NP}{\overline{\text{sets}}} \quad \cfrac{\underset{N/N}{\overline{\text{new}}} \quad \underset{N}{\overline{\text{record}}}}{\cfrac{N}{NP}}}{S\backslash NP} \\ \overline{ S }$$

This example exhibits the *Subject-Verb-Object* construction typical of English. Here we see the verb acting as a function between the subject and object, and uses the following rule: **(S \NP)/NP**. To evaluate this, a given noun phrase should exist to the right of the function (in this case the verb) to produce the function **(S \NP)**. This then evaluates to give a sentence when a noun phrase exists to its left. If we take the phrase *'new record'*, the adjective new has the complex type **N/N**, which merely acts as a recursive function. These compositional functions enable the valence shifters described in the previous subsection to be integrated into our approach to emotion classification.

The derivation can be described through use of the semantic structure shown in Figure 1.

The System

Our system integrates four modules to tackle the problem of emotion classification. These are: an augmented version of WN-A, which takes into account emotion bearing concepts which may have been present in headlines at the time of the task, a text-normalization unit, a CCG parser (Clark and Curran 2004), and a lexical lookup module, dependent on the output of the contextual valence shifters, which is used to determine whether an emotional term appears in the valence-classified headline. The valence shifters that we used were adapted versions of those presented in Simančík and Lee (2009) and Polanyi and Zaenen (2006).

Extension of WordNet-Affect

Emotion	Associated Concepts
Anger	seize, war, bomb, sanction, attack
Disgust	porn, kidnap, desecrate, violence
Fear	Iraq, Gaza, cancer, massacre, terror, Al Qaeda
Joy	win, fun, pleasure, celebrate
Sadness	misfortune, cancel, kill, widow
Surprise	realise, discover, shock

Table 3: Emotional concept bearing words

The version of WordNet-Affect (WN-A) provided by the Affective Task organisers contained a set of emotion-bearing unigrams associated with the six relevant categories of the headline corpus. The terms included in this lexicon are general terms for describing an emotion, and would be useful in cross-domain classification, where the communication of emotion in text is explicit. Strapparava, Valitutti, and Stock (2006) would describe these terms in the lexicon

	Accuracy	Precision	Recall	F1
Anger	97.82	28.57	10.53	15.38
Disgust	99.11	66.67	41.67	47.70
Fear	90.74	43.75	15.73	23.14
Joy	88.44	39.13	16.98	23.68
Sadness	90.93	57.15	32.32	41.29
Surprise	93.20	20.83	25.00	22.72
System Average	**93.35**	**42.68**	**23.70**	**28.97**
UPAR7 Comparison	89.43	27.61	5.69	8.71

Table 2: Results from our final system

as direct affective words. Nevertheless, the corpus involved contained headlines, which were mostly less than ten words in length, and contained few of the explicit emotion-bearing terms. Due to the implicit nature of emotional expression in the headlines, it became clear through a qualitative analysis of the training set that emotions were being associated with specific concepts and events that were the subject of the headlines.

We compiled a list of emotion bearing concepts based upon the training set and related ideas, that we believed would be pertinent within the genre of news story headlines for the period of time when the corpus was compiled, 2007. Table 3 outlines some of the lexical items that we initially compiled.

In order to augment these initial concepts we used Word-Net 3.0 (Fellbaum 1998). For the adjectives we explored and added any unique terms discovered via the *similar to* links, which helped maintain the original meaning of our set of seeds. For the nouns and verbs in the seed set we explored the hyponymic links to extend our seed set.

Results

Table 2 shows the results from experimentation with our system on the test dataset, consisting of 1,000 headlines. Over the six emotional categories, our system achieved an average accuracy of 93.35%, an increase of 3.92% over the previous best system for the task, UPAR7 (Chaumartin 2007). In the remaining coarse-grained metrics, our system also outperformed the previous best system. Our system average for precision was 42.68% , an increase of 15.07% , and our average recall value was 23.70% , also yielding a gain of 18.01% . Our resulting F1 measure delivered an increase of 20.26% .

If we consider the results on the emotion categories themselves, our system also performed favourably. In particular, the category of disgust performed well across all metrics, with a resulting accuracy of 99.11% and an F1 score of 47.70% . This can be attributed to the relatively small number of headlines labelled with the category of disgust in the test set (1.2%), which seem to describe similar news stories (such as *porn*).

Sadness also yields good results. Whilst only achieving a recall value of 32.32% , the precision sits at 57.15%, which is above the random baseline, even for a polarity based sentiment classification task. Fear and joy also share high precision values, at 43.75% and 39.13% respectively.

The classes of emotion that did not yield comparable results to the other emotional classes that were categorised during experimentation were *Anger* and *Surprise*. Anger yielded the lowest value of recall, at 10.53% and surprise the lowest precision score, at 20.83% .

Related Work

This section will highlight some of the systems for Sentiment Analysis that have been developed specifically for use with the headline corpus.

Systems Developed for the Emotion Classification Task

Several systems participated in the SemEval Task 14 emotion classification task. UPAR 7, a system developed by Chaumartin (2007), delivered the best performance on the emotion classification task. UPAR7 utilised an enriched version of SentiWordNet (Esuli and Sebastiani 2006) and WordNetAffect as the base lexica for the task. Alongside these resources, the Stanford parser was used to identify salient head word structures in the headlines, and valence shifting rules based on the work of Polanyi and Zaenen (2006) were additionally implemented. The system bears a resemblance to our approach, and their final rule-based system yielded an average accuracy of 89.43% over the six-emotions of the task.

The SWAT system, developed by Katz, Singleton, and Wicentowski (2007), expand their training set to include an additional 1,000 headlines from the Associated Press. These were duly annotated by non-expert, untrained annotators. Roget's New Millennium Thesaurus is used to create an extensive word to emotion mapping, and this is used as SWAT's lexicon. The average accuracy achieved by the system was 88.58%, and is ranked second out of the participating systems.

The final system to take part in the emotion classification task was the UA system, developed by Kozareva et al. (2007). Their system approaches emotion classification by observing word-frequency and co-occurrence counts within online documents. They base this on the hypothesis that words which co-occur across a document-set annotated with a given emotion exhibit a high probability of expressing a particular emotion. Kozareva et al. (2007) note that they do not consider the impact of valence shifters in their work, and the shifting roles that adverbs and adjectives perform, and this may possibly have affected their overall performance.

The system returns an average accuracy of 85.72% over the test set. Full results for the participating system are shown in Table 4 .

		Accuracy	Precision	Recall	F1
	Anger				
SWAT	24.51	92.10	12.00	5.00	7.06
UA	23.20	86.40	12.74	**21.6**	**16.03**
UPAR7	**32.33**	**93.60**	**16.67**	1.66	3.02
	Disgust				
SWAT	**18.55**	97.20	0.00	0.00	-
UA	16.21	**97.30**	0.00	0.00	-
UPAR7	12.85	95.30	0.00	0.00	-
	Fear				
SWAT	32.52	84.80	25.00	14.40	18.27
UA	23.15	75.30	16.23	**26.27**	**20.06**
UPAR7	**44.92**	**87.90**	**33.33**	2.54	4.72
	Joy				
SWAT	**26.11**	80.60	35.41	**9.44**	**14.91**
UA	2.35	81.80	40.00	2.22	4.21
UPAR7	22.49	**82.20**	**54.54**	6.66	11.87
	Sadness				
SWAT	38.98	87.70	32.50	11.92	17.44
UA	12.28	88.90	25.00	0.91	1.76
UPAR7	**40.98**	**89.00**	**48.97**	**22.02**	**30.38**
	Surprise				
SWAT	11.82	**89.10**	11.86	10.93	11.78
UA	7.75	84.60	**13.70**	**16.56**	**15.00**
UPAR7	**16.71**	88.60	12.12	1.25	2.27

Table 4: System results from the emotion classification task (Strapparava and Mihalcea 2007)

Other systems utilising the Headline Corpus A number of other systems developed for emotion classification post-competition also use the headline corpus as a test set for their algorithms. Mohammad (2012) created six binary classifiers for the emotions present in the headline corpus, and experimented with Logistic Regression and Support Vector Machines approaches. As supervised learning methods require sufficient data to perform adequately, the experiments deviated from the scope of the SemEval Affective Task, which was to create an emotion classification system in an unsupervised environment. The system performs well when the roles of training and test sets are swapped, but the role of training set size in overall performance should be considered. Kirange and Deshmukh (2012) also approach the task with a similar Support Vector Machines based system.

Discussion

In the following section we will discuss the following points in regards to our results:

- The effects of contextual valence shifters
- The inherent subjectivity associated with annotating emotions
- The role of surprise within the emotion classification spectrum

Effects of Contextual Valence Shifters

To discuss the effect that contextual valence shifters have on the task of emotion classification of headlines, it will be worth comparing our system to a basic lexical matching system, with no rules or stipulations, that uses the WordNet-Affect lexicon. The results of this are shown in Table 5.

	Accuracy	Precision	Recall	F1
Anger	97.70	25.00	10.53	14.81
Disgust	98.67	0	0	0
Fear	91.22	52.00	14.61	22.81
Joy	82.42	11.96	10.37	11.11
Sadness	89.20	26.31	5.05	8.47
Surprise	94.90	13.33	5.00	7.27

Table 5: Results from using WN-A only

If we compare the accuracy scores, improvements are only slight. However, we must remember that accuracy also takes into account false positives when calculating the overall results. If we combine this with the fact that when removing annotation scores of lower than 50 to carry out the coarse-grained evaluation of our system, then we discover that 66.5% of the headlines are classed as emotionless in the test set, despite their salience in fact being minimal. Neutral instances in sentiment classification always pose a problem, and we believe that our system deals with these appropriately, as can be seen from the gains in precision and recall over a basic lexical matching approach.

The attribute that we believe has given considerable strength to our method is the assumption that emotions are valenced. We attribute the results in general to the integration of contextual valence shifters to our system. The work of Simančík and Lee (2009) demonstrated the effectiveness of contextual valence shifters on the task in hand, and by incorporating this approach into our system, we believe that this produced the relevant increases in accuracy, precision and recall.

Interestingly enough also, UPAR 7 (Chaumartin 2007), the previously best performing system on the emotion classification task, also utilised valence shifters in their work, which produced favourable results in comparison to the other systems. What their system may have lacked however, is the combination with a suitable grammar, such as CCG, in order to access the compositional semantics of the headlines being classified.

Comparison with Inter-Annotator Agreement

The results from our system may compare favourably to other unsupervised systems proposed for the task, but irrespective of this, our results are not exceptionally high. While this may make the system appear weak, the difficulty with recognising emotions amongst humans must be introduced, so as to give some context to the achievements of our system. Six annotators were asked to annotate the dataset with the six proposed emotions, and the results of evaluating these annotations using the Pearson correlation method are shown in Table 6.

As can be seen here, the levels of agreement do not go above 70%, and the emotion with the highest agreement is *Sadness*, at an average level of 68.19% agreement. This highlights the difficulties of annotating emotion, due to their highly subjective nature. This leads to varying levels of disagreement amongst the annotators. One particular emotion which annotators struggled to agree on is that of *surprise*.

Emotion	Agreement Score
Anger	49.55
Disgust	44.51
Fear	63.81
Joy	59.91
Sadness	68.19
Surprise	36.07

Table 6: Inter-annotator agreement scores (Strapparava and Mihalcea 2007)

The Element of Surprise

The one emotion which both our system and others that participated in the Affective Task struggle to classify with satisfactory precision and recall is surprise. Despite outperforming other systems, with ours achieving a precision of 20.83% and recall of 25.00% , these figures are still relatively low in comparison with the other categories.

The inclusion of surprise in any corpus of emotion-bearing text is an interesting one, and may be attributed to the work of Ekman (1982). However it sits in a different zone to the other emotions that are discussed throughout the task. If we refer to the work of Ortony, Clore, and Collins (1988) once again, they struggle to class surprise as an emotion, due to the inherently neutral nature which it can adopt. This facet is mirrored in the headlines which were annotated as containing strong elements of surprise in the headline corpus. Quite often, seemingly neutral lexical items in headlines such as *discovery* flag that a headline conveys a form of surprise. This leads to difficulties in compiling a lexicon of emotional terms related to surprise, as generally, explicit items will form the majority of this lexicon. Careful consideration of the domain, and observing token terms that are not necessarily emotion bearing, is what helped to produce the classification results for this category in our system. Due to the inherent difficulties outlined with this particular category of emotion, further corpus analysis of this phenomena is required, in particular focussing on the lexical entities associated with this emotion across domains.

Conclusions

We have developed a system for the classification of emotions held in headlines, which yields favourable classifications results in comparison with other similar systems. For the headlines in 2007 SemEval Affective Task emotion-labelled test set, our system produced higher accuracy, precision and recall scores on average than the top performing systems. The integration of the CCG parser to yield each headline's underlying compositional semantics in combination with the contextual valence shifters seems to be a very promising combination for automatic emotion annotation in headlines. To improve the scores further, an in depth understanding of the context of the domain could be integrated with the lexicon. The category of surprise also requires further study, as the available literature seems limited, yet implementing a suitable system could have positive effects on the study of automatic emotion classification. Supervised approaches to emotion classification, such as the work of Mohammad (2012) yields fruitful results, and if contextual valence shifters were integrated with this, it is believed that further increases in classification precision and recall could be produced.

Our system highlights the importance of contextual valence shifting when approaching emotion labelling. Through this work, and the successful work of others (Chaumartin 2007; Polanyi and Zaenen 2006) we argue that compositional semantic based valence shifters are a vital part of any system undertaking semi-supervised sentiment analysis, under the assumption that emotions are valence-rooted.

References

Blitzer, J.; Dredze, M.; and Pereira, F. 2007. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In *Annual Meeting-Association For Computational Linguistics*, volume 45, 440.

Chaumartin, F. 2007. UPAR7: A knowledge-based system for headline sentiment tagging. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, 422–425. Prague, Czech Republic: Association for Computational Linguistics.

Clark, S., and Curran, J. R. 2004. Parsing the wsj using ccg and log-linear models. In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04)*, 104–111.

Ekman, P. 1982. *Emotion in the human face*. New York: Cambridge University Press.

Esuli, A., and Sebastiani, F. 2006. SentiWordNet: A publicly available lexical resource for opinion mining. In *Proceedings of LREC*, volume 6, 417–422.

Fellbaum, C. 1998. *WordNet: An Electronic Lexical Database*. Cambridge, MA, USA: MIT Press.

Hansen, L. K.; Arvidsson, A.; rup Nielsen, F.; Colleoni, E.; and Etter, M. 2011. Good Friends, Bad News - Affect and Virality in Twitter. In *The 2011 International Workshop on Social Computing, Network, and Services (SocialComNet 2011)*.

Katz, P.; Singleton, M.; and Wicentowski, R. 2007. SWAT-MP:The SemEval-2007 Systems for Task 5 and Task 14. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, 308–313. Prague, Czech Republic: Association for Computational Linguistics.

Kirange, D., and Deshmukh, R. 2012. Emotion Classification of News Headlines Using SVM. In *Asian Journal of Computer Science and Information Technology*, 104–106.

Kozareva, Z.; Navarro, B.; Vazquez, S.; and Montoyo, A. 2007. UA-ZBSA: A Headline Emotion Classification through Web Information. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, 334–337. Prague, Czech Republic: Association for Computational Linguistics.

Mohammad, S. 2012. Portable features for classifying emotional text. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 587–591. Montréal, Canada: Association for Computational Linguistics.

Mukherjee, A., and Liu, B. 2012. Mining contentions from discussions and debates. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '12, 841–849. New York, NY, USA: ACM.

Ortony, A.; Clore, G. L.; and Collins, A. 1988. *The Cognitive Structure of Emotions*. New York: Cambridge University Press.

Polanyi, L., and Zaenen, A. 2006. Contextual valence shifters. *Computing attitude and affect in text Theory and applications* 1–10.

Simančík, F., and Lee, M. 2009. A CCG-based system for valence shifting for sentiment analysis. *Research in Computing Science* 41:99–108.

Smith, P., and Lee, M. 2012. Cross-discourse Development of Supervised Sentiment Analysis in the Clinical Domain. In *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis*, 79–83. Jeju, Korea: Association for Computational Linguistics.

Steedman, M. 2000. *The Syntactic Process*. Cambridge, MA, USA: The MIT Press.

Strapparava, C., and Mihalcea, R. 2007. SemEval-2007 Task 14: Affective Text. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, 70–74. Prague, Czech Republic: Association for Computational Linguistics.

Strapparava, C., and Valitutti, A. 2004. WordNet-Affect: an Affective Extension of WordNet. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*.

Strapparava, C.; Valitutti, A.; and Stock, O. 2006. The affective weight of lexicon. In *Proceedings of 5th International Conference on Language Resources and Evaluation (LREC 2006)*, 423–426.