In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)

Learning Emotion Indicators from Tweets: Hashtags, Hashtag Patterns, and Phrases

Ashequl Qadir
School of Computing
University of Utah
Salt Lake City, UT 84112, USA
asheq@cs.utah.edu

Ellen Riloff
School of Computing
University of Utah
Salt Lake City, UT 84112, USA
riloff@cs.utah.edu

Abstract

We present a weakly supervised approach for learning hashtags, hashtag patterns, and phrases associated with five emotions: AFFECTION, ANGER/RAGE, FEAR/ANXIETY, JOY, and SADNESS/DISAPPOINTMENT. Starting with seed hashtags to label an initial set of tweets, we train emotion classifiers and use them to learn new emotion hashtags and hashtag patterns. This process then repeats in a bootstrapping framework. Emotion phrases are also extracted from the learned hashtags and used to create phrase-based emotion classifiers. We show that the learned set of emotion indicators yields a substantial improvement in F-scores, ranging from +%5 to +%18 over baseline classifiers.

1 Introduction

Identifying emotions in social media text can be beneficial for many applications, for example to help companies understand how people feel about their products, to assist governments in recognizing growing anger or fear associated with an event, or to help media outlets understand people's emotional response toward controversial issues or international affairs. On the Twitter micro-blogging platform, people often use hashtags to express an emotional state (e.g., *#happyasalways*, *#angryattheworld*). While some hashtags consist of a single word (e.g., *#angry*), many hashtags include multiple words and creative spellings (e.g., *#cantwait4tmrw, #Youredabest*), which can not be easily recognized using sentiment or emotion lexicons.

Our research learns three types of emotion indicators for tweets: hashtags, hashtag patterns, and phrases for one of five emotions: AFFECTION, ANGER/RAGE, FEAR/ANXIETY, JOY, or SADNESS/DISAPPOINTMENT. We present a bootstrapping framework for learning emotion hashtags and extend the framework to also learn more general hashtag patterns. We then harvest emotion phrases from the hashtags and hashtag patterns for contextual emotion classification.

First, we make the observation that emotion hashtags often share a common prefix. For example, *#angryattheworld* and *#angryatlife* both have the prefix *"angry at"*, which suggests the emotion ANGER. Consequently, we generalize beyond specific hashtags to create *hashtag patterns* that will match all hashtags with the same prefix, such as the pattern *#angryat** which will match both *#angryattheworld* and *#angryatlife*.

A key challenge is that a seemingly strong emotion word or phrase can have a different meaning depending upon the following words. For example, *#angry** may seem like an obvious pattern to identify ANGER tweets. But *#angrybirds* is a popular hashtag that refers to a game, not the writer's emotion. Similarly, "love you" usually expresses AFFECTION when it is followed by a person (e.g., *#loveyoumom*). But it may express JOY in other contexts (e.g., *#loveyoulife*). We use probability estimates to determine which hashtag patterns are reliable indicators for an emotion.

Our second observation is that hashtags can also be used to harvest emotion phrases. For example, if we learn that the hashtag *#lovelife* is associated with JOY, then we can extract the phrase "love life" from the hashtag and use it to recognize emotion in the body of tweets. However, unlike hashtags, which are self-contained, the words surrounding a phrase in a tweet must also be considered. For example, negation can toggle polarity (*"don't love life"* may suggest SADNESS, not JOY) and the aspectual context may indicate that no emotion is being expressed (e.g., *"I would love life if ..."*). Consequently, we train classifiers to determine if a tweet contains an emotion based on both an emotion phrase and its context.

2 Related Work

In addition to sentiment analysis, which has been widely studied (e.g., (Barbosa and Feng, 2010; Brody and Diakopoulos, 2011; Kouloumpis et al., 2011; Mitchell et al., 2013)), recognizing emotions in social media text has also become a popular research topic in recent years. Researchers have studied feature sets and linguistic styles (Roberts et al., 2012), emotion influencing behaviors (Kim et al., 2012), sentence contexts (Yang et al., 2007b), hierarchical emotion classification (Ghazi et al., 2010; Esmin et al., 2012) and emotion lexicon creation (Yang et al., 2007a; Mohammad, 2012a; Staiano and Guerini, 2014). Researchers have also started to utilize the hashtags of tweets, but primarily to collect labeled data (e.g., for sarcasm (Davi-

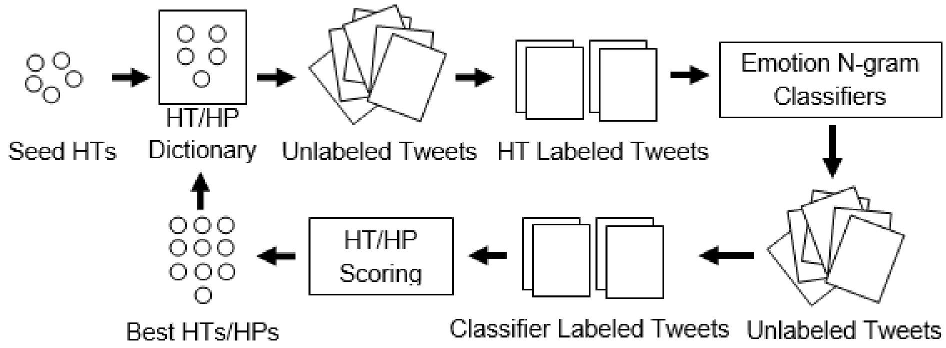

Figure 1: Bootstrapped Learning. (HT = hashtag; HP = hashtag pattern)

dov et al., 2010; Riloff et al., 2013) and for sentiment/emotion data (Wang et al., 2012; Mohammad et al., 2013; Choudhury et al., 2012; Purver and Battersby, 2012; Mohammad, 2012a)).

Wang et al. (2011) investigated several graph based algorithms to collectively classify hashtag sentiments, but their work is focused on positive versus negative polarity classification. Our research extends the preliminary work on bootstrapped learning of emotion hashtags (Qadir and Riloff, 2013) to additionally learn patterns corresponding to hashtag prefix expressions and to extract emotion phrases from the hashtags, which are used to train phrase-based emotion classifiers.

Emotion Classes	**Seed Hashtags**
AFFECTION	*#loveyou*, *#sweetheart*, *#bff* *#romantic*, *#soulmate*
ANGER & RAGE	*#angry*, *#mad*, *#hateyou* *#pissedoff*, *#furious*
FEAR & ANXIETY	*#afraid*, *#petrified*, *#scared* *#anxious*, *#worried*
JOY	*#happy*, *#excited*, *#yay* *#blessed*, *#thrilled*
SADNESS & DISAPPOINTMENT	*#sad*, *#depressed* *#disappointed*, *#unhappy* *#foreveralone*

Table 1: Emotion Classes and Seed Hashtags

3 Learning Emotion Hashtags, Hashtag Patterns and Phrases

For our research, we collapsed Parrot's emotion taxonomy (Parrott, 2001)[1] into 5 emotion classes that frequently occur in tweets and minimally overlap with each other: AFFECTION, ANGER/RAGE, FEAR/ANXIETY, JOY, and SADNESS/DISAPPOINTMENT. We also used a NONE OF THE ABOVE class for tweets that do not express any emotion or express an emotion different from our five classes. For each of these categories, we identified 5 common hashtags that are strongly associated with the emotion and used them as seeds. Table 1 shows the seed hashtags.

Compared to the Ekman emotion classes (Ekman, 1992), one of the emotion taxonomies frequently used in NLP research (Strapparava and Mihalcea, 2007; Mohammad, 2012b), JOY, ANGER, SADNESS and FEAR are comparable to 4 of our 5 emotion classes. We do not study Ekman's SURPRISE and DISGUST classes, but include AFFECTION.

3.1 Learning Hashtags

Figure 1 presents the framework of the bootstrapping algorithm for hashtag learning. The process begins by collecting tweets that contain the seed hashtags and labeling them with the corresponding emotion. For this purpose, we collected 323,000 tweets in total that contain at least one of our seed hashtags. We also exploit a large pool of *unlabeled tweets* to use during bootstrapping, consisting of 2.3 million tweets with at least one hashtag per tweet (because we want to learn hashtags), collected using Twitter's streaming API. We did not include retweets or tweets with URLs, to reduce duplication and focus on tweets with original content. The *unlabeled tweets* dataset had 1.29 average hashtags-per-tweet and 3.95 average tweets-per-hashtag. We preprocessed the tweets with CMU's tokenizer (Owoputi et al., 2013) and normalized with respect to case.

The labeled tweets are then used to train a set of emotion classifiers. We trained one logistic regression classifier for each emotion class using the LIBLINEAR package (Fan et al., 2008). We chose logistic regression because it produces probabilities with its predictions, which are used to assign scores to hashtags. As features, we used unigrams and bigrams with frequency $>$ 1. We removed the seed hashtags from the tweets so the classifiers could not use them as features.

For each emotion class $e \in E$, the tweets containing a seed hashtag for e were used as positive training instances. The negative training instances consisted of the tweets containing seed hashtags for the competing emotions as well as 100,000 randomly selected tweets

[1]There were other emotions in Parrott's taxonomy such as SURPRISE, NEGLECT, etc. that we did not use for this research.

Affection	Anger & Rage	Fear & Anxiety	Joy	Sadness & Disappointment
#yourthebest	#godie	#hatespiders	#tripleblessed	#leftout
#myotherhalf	#donttalktome	#haunted	#tgfad	#foreverugly
#bestfriendforever	#pieceofshit	#shittingmyself	#greatmood	#singleprobs
#loveyoulots	#irritated	#worstfear	#thankful	#lonerlyfe
#flyhigh	#fuming	#scaresme	#atlast	#teamlonely
#comehomesoon	#hateliars	#nightmares	#feelinggood	#unloved
#wuvyou	#heated	#paranoid	#happygirl	#friendless
#alwaysandforever	#getoutofmylife	#hateneedles	#godisgreat	#heartbroken
#missyousomuch	#angrytweet	#frightened	#superhappy	#needalife
#loveyougirl	#dontbothermewhen	#freakedout	#ecstatic	#letdown

Table 2: Examples of Learned Hashtags

from our unlabeled tweets. Although some of the unlabeled tweets may correspond to emotion e, we expect that most will have no emotion or an emotion different from e, giving us a slightly noisy but large, diverse set of negative instances.

We then apply each emotion classifier to the unlabeled tweets. For each emotion e, we collect the tweets classified as e and extract the hashtags from those tweets to create a candidate pool H_e of hashtags for emotion e. To limit the number of candidates, we discard hashtags that occur < 10 times, have just one character, or have > 20 characters. Next, we score each candidate hashtag h by computing the average probability assigned by the logistic regression classifier for emotion e over all of the tweets containing hashtag h. For each emotion class, we select the 10 hashtags with the highest scores. From the *unlabeled tweets*, we then add all tweets with one of the learned hashtags to the training instances, and the bootstrapping process continues. Table 2 shows examples of the learned hashtags.

3.2 Learning Hashtag Patterns

We learn hashtag patterns in a similar but separate bootstrapping process. We first expand each hashtag into a sequence of words using an N-gram based word segmentation algorithm[2] supplied with corpus statistics from our tweet collection. For example, *#angryatlife* expands[3] to the phrase *"angry at life"*. We use a Prefix Tree (Trie) data structure to represent all possible prefixes of the expanded hashtag phrases, but the prefixes consist of words instead of characters.

Next, we traverse the tries and consider all possible prefix paths as candidate hashtag patterns. We only consider prefixes that have occurred with at least one following word. For example, *#angryashell*, *#angryasalways*, *#angrybird*, *#angryatlife*, *#angryatyou* would produce patterns: *#angry**, *#angryas**, *#angryat** as shown in Figure 2.

We score each pattern by applying the classifier for

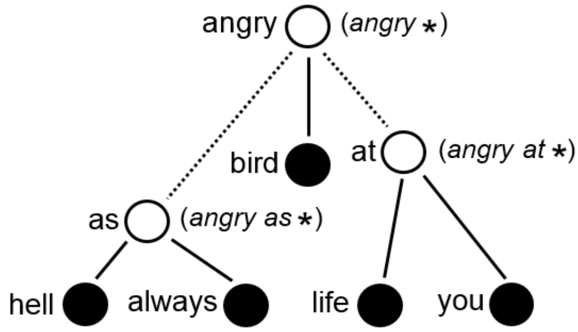

Figure 2: Trie of example hashtags with prefix *angry*. Dotted lines lead to non-terminal nodes where patterns are extracted.

emotion e (trained in the same way as hashtag learning) to all tweets having hashtags that match the pattern. We compute the average probability produced by the classifier, and for each emotion class, we select the 10 hashtag patterns with the highest scores. From the *unlabeled tweets*, we then add all tweets with hashtags that match one of the learned hashtag patterns to the training instances, and the bootstrapping process continues. Table 3 shows examples of learned hashtag patterns and matched hashtags.

3.3 Creating Phrase-based Classifiers

The third type of emotion indicator that we acquire are emotion phrases. At the end of the bootstrapping process, we apply the word segmentation algorithm to all of the learned hashtags and hashtag patterns to expand them into phrases (e.g., *#lovemylife* $\rightarrow$ *"love my life"*). Each phrase is assumed to express the same emotion as the original hashtag. However, as we will see in Section 4, just the presence of a phrase yields low precision, and surrounding context must also be taken into account.

Consequently, we train a logistic regression classifier for each emotion e, which classifies a tweet with respect to emotion e based on the presence of a learned phrase for e as well as a context window of size 6 around the phrase (set of 3 words on its left and set of 3

[2] http://norvig.com/ngrams/

[3] On a random sample of 100 hashtags, we found expansion accuracy to be 76% (+8% partially correct expansions).

Emotion	Hashtag Pattern	Examples of Matching Hashtags
AFFECTION	#bestie*	#bestiefolyfe, #bestienight, #bestielove
	#missedyou*	#missedyoutoomuch, #missedyouguys, #missedyoubabies
ANGER & RAGE	#godie*	#godieoldman, #godieyou, #godieinahole
	#pissedoff*	#pissedofffather, #pissedoffnow, #pissedoffmood
FEAR & ANXIETY	#tooscared*	#tooscaredtogoalone, #tooscaredformama, #tooscaredtomove
	#nightmares*	#nightmaresfordays, #nightmaresforlife, #nightmarestonight
JOY	#feelinggood*	#feelinggoodnow, #feelinggoodforme, #feelinggoodabout
	#goodmood*	#goodmooditsgameday, #goodmoodmode, #goodmoodnight
SADNESS & DISAPPOINTMENT	#bummed*	#bummedout, #bummedaf, #bummednow
	#singlelife*	#singlelifeblows, #singlelifeforme, #singlelifesucks

Table 3: Examples of Learned Hashtag Patterns and Matching Hashtags

words on its right). Tweets containing a learned phrase for e and a seed hashtag for e are the positive training instances. Tweets containing a learned phrase for e and a seed hashtag for a different emotion are used as the negative training instances. For example, when *"love my life"* is learned as an emotion phrase for JOY, the tweet, *"how can I love my life when everybody leaves me! #sad"* will have one feature each for the left words *"how"*, *"can"*, and *"I"*, one feature each for the right words *"when"*, *"everybody"* and *"leaves"*, and one feature for the phrase *"love my life"*. The tweet will then be considered a negative instance for JOY because *"#sad"* indicates a different emotion.

4 Experimental Results

To evaluate our learned emotion indicators, we manually selected 25 topic keywords/phrases[4] that we considered to be strongly associated with emotions, but not necessarily with any specific emotions of our study. We then searched in Twitter using Twitter Search API for any of these topic phrases and their corresponding hashtags. These 25 topic phrases are: *Prom, Exam, Graduation, Marriage, Divorce, Husband, Wife, Boyfriend, Girlfriend, Job, Hire, Laid Off, Retirement, Win, Lose, Accident, Failure, Success, Spider, Loud Noise, Chest Pain, Storm, Home Alone, No Sleep* and *Interview*. Since the purpose is to evaluate the quality and coverage of the emotion hashtags that we learn, we filtered out any tweet that did not have at least one hashtag.

Two annotators were given annotation guidelines and were instructed to label each tweet with up to two emotions. The instructions specified that the emotion must be felt by the writer. The annotators annotated 500 tweets with an inter-annotator agreement level of 0.79 Kappa (κ) (Carletta, 1996). The annotation disagreements in these 500 tweets were then adjudicated, and each annotator labeled an additional 2,500 tweets. Altogether this gave us an emotion annotated dataset of 5,500 tweets. We randomly separated out 1,000 tweets from this collection as a tuning set, and used the remaining 4,500 tweets as evaluation data. The distribution of emotions in the evaluation data was 6% for AFFECTION, 9% for ANGER/RAGE, 13% for FEAR/ANXIETY, 22% for JOY, and 12% for SADNESS/DISAPPOINTMENT. 42% of the tweets had none of the 5 emotions and 4% of the tweets had more than one emotions in the same tweet.

We created two baseline systems to assess the difficulty of the emotion classification task. First, we created SVM classifiers for each emotion using N-gram features and performed 10-fold cross-validation on the test data. We used LIBSVM (Chang and Lin, 2011) and set the $cost$ and $gamma$ parameters based on the tuning data. Second, we acquired the NRC Emotional Tweets Lexicon (Mohammad, 2012a), which contains emotion unigrams and bigrams for 8 emotions, 4 that are comparable to ours: ANGER, FEAR, JOY and SADNESS. We created a hashtag from each term in the lexicon by appending a # symbol on the front and removing whitespace. For each term, we chose the emotion with the highest score in the lexicon.

Table 4 shows our experimental results. The baseline classifiers (SVM_1 uses unigrams, SVM_{1+2} uses unigrams and bigrams) have low recall but 63-78% precision. The hashtags created from the NRC Lexicon have low precision. This could be due to possible entries (e.g., *"candy"* or *"idea"*), which without context are not much indicative of any specific emotion.

The second section of Table 4 shows the results when we label a tweet based on the presence of a hashtag or hashtag pattern. First, we use just the 5 seed hashtags to assess their coverage (as expected, high precision but low recall). Next, we add the hashtags learned during bootstrapping. For most emotions, the hashtags achieve performance similar to the supervised SVMs. The following row shows results for our learned hashtag patterns. Recall improves by +14% for AFFECTION, which illustrates the benefit of more general hashtag patterns, and at least maintains similar level of precision for other emotions. When the hashtags and hashtag patterns are combined (HTs+HPs), we see the best of both worlds with improved recall as high as +17% in AFFECTION and +10% in FEAR/ANXIETY

[4]This data collection process is similar to the emotion tweet dataset creation by Roberts et al. (2012)

Method	AFFECTION	ANGER & RAGE	FEAR & ANXIETY	JOY	SADNESS & DISAPPOINT.
	P R F	P R F	P R F	P R F	P R F
Baselines					
SVM_1	78 40 53	66 17 27	68 33 44	66 47 55	63 26 37
SVM_{1+2}	78 35 48	67 10 17	68 29 41	65 43 52	63 21 32
NRC Lexicon HTs	n/a	26 16 20	39 12 18	36 13 19	28 18 22
Learned Hashtags (HTs) and Hashtag Patterns (HPs)					
Seed HTs	94 06 11	75 01 03	100 06 11	93 04 08	81 02 05
All HTs	82 34 48	63 23 34	60 37 46	81 13 22	72 28 40
All HPs	76 48 59	60 22 32	57 42 48	84 09 16	73 16 26
All HTs+HPs	74 51 60	56 27 36	55 47 51	80 15 25	70 29 41
Learned Emotion Phrases					
Emotion Phrases	32 28 30	17 46 25	28 45 35	50 23 32	26 30 28
Phrase-based Classifier (PC)	54 07 12	48 05 09	63 17 27	69 12 20	50 06 11
SVM_1+PC	79 42 55	63 18 28	70 35 47	68 48 56	62 27 38
Hybrid Approach					
SVM_1+PC ∪ HTs+HPs	69 64 **66**	55 38 **45**	54 61 **57**	68 54 **60**	62 44 **51**

Table 4: Emotion Classification Results (P = Precision, R = Recall, F = F-score)

compared to All HTs, as well as improved F-scores across the board.

The third section of Table 4 presents the results for the emotion phrases. The first row (Emotion Phrases) shows that labeling a tweet based solely on the presence of a phrase is not very accurate. Next, we applied the trained models of the phrase-based classifiers (described in Section 3.3) to each tweet of the evaluation data. This provided us with probability of an emotion for each of the 5 emotions. The phrase-based classifiers (PC) yield higher precision, albeit with low recall. Finally, we use these probabilities as 5 additional features to SVM_1. The corresponding SVM_1+PC row shows a consistent 1-2 point F score gain over the original SVM_1 baseline.

The last section of Table 4 shows the best results with a hybrid system, which labels a tweet with emotion e if EITHER the enhanced SVM labels it as e OR the tweet contains a hashtag or hashtag pattern associated with e. This combined approach achieves substantially higher performance than any individual method across all 5 emotion classes, with improved F-scores ranging from +%5 to +%18 over the baseline classifiers, demonstrating that the different types of emotion indicators are complementary.

5 Conclusions

We have shown that three types of emotion indicators can be learned from tweets with weakly supervised bootstrapping: hashtags, hashtag patterns, and phrases. Our findings suggest that emotion hashtags are strong indicators for recognizing writer's emotion in tweets, and can be further generalized into hashtag patterns by learning prefix expressions corresponding to an emotion. Phrases learned from the hashtags and patterns are not always reliable by themselves, but training additional classifiers with the emotion phrases and their surrounding context provides added benefits to emotion classification in tweets. Our results showed that combining the learned emotion indicators with an N-gram classifier in a hybrid approach substantially improves performance across 5 emotion classes.

Acknowledgments

This work was supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center (DoI/NBC) contract number D12PC00285. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the U.S. Government.

References

Luciano Barbosa and Junlan Feng. 2010. Robust sentiment detection on twitter from biased and noisy data. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, COLING '10.

Samuel Brody and Nicholas Diakopoulos. 2011. Cooooooooooooooollllllllllllll!!!!!!!!!!!!!!: using word lengthening to detect sentiment in microblogs. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '11.

Jean Carletta. 1996. Assessing agreement on classification tasks: the kappa statistic. *Comput. Linguist.*, 22:249–254, June.

Chih-Chung Chang and Chih-Jen Lin. 2011. Libsvm: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3):27:1–27:27, May.

Munmun De Choudhury, Michael Gamon, and Scott Counts. 2012. Happy, nervous or surprised? classification of human affective states in social media. In *Proceedings of the Sixth International Conference on Weblogs and Social Media*.

Dmitry Davidov, Oren Tsur, and Ari Rappoport. 2010. Semi-supervised recognition of sarcastic sentences in twitter and amazon. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning*, CoNLL '10.

Paul Ekman. 1992. An argument for basic emotions. *Cognition and Emotion*, 6(3):169200.

Ahmed Ali Abdalla Esmin, Roberto L. De Oliveira Jr., and Stan Matwin. 2012. Hierarchical classification approach to emotion recognition in twitter. In *Proceedings of the 11th International Conference on Machine Learning and Applications, ICMLA, Boca Raton, FL, USA, December 12-15, 2012. Volume 2*, pages 381–385. IEEE.

Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. 2008. Liblinear: A library for large linear classification. *J. Mach. Learn. Res.*, 9:1871–1874, June.

Diman Ghazi, Diana Inkpen, and Stan Szpakowicz. 2010. Hierarchical versus flat classification of emotions in text. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, CAAGET '10.

Suin Kim, JinYeong Bak, and Alice Oh. 2012. Discovering emotion influence patterns in online social network conversations. *SIGWEB Newsl.*, (Autumn):3:1–3:6, September.

Efthymios Kouloumpis, Theresa Wilson, and Johanna Moore. 2011. Twitter sentiment analysis: The good the bad and the omg! In *Proceedings of the Fifth International Conference on Weblogs and Social Media*.

Margaret Mitchell, Jacqui Aguilar, Theresa Wilson, and Benjamin Van Durme. 2013. Open domain targeted sentiment. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*.

Saif Mohammad, Svetlana Kiritchenko, and Xiaodan Zhu. 2013. Nrc-canada: Building the state-of-the-art in sentiment analysis of tweets. In *Proceedings of the seventh international workshop on Semantic Evaluation Exercises (SemEval-2013)*.

Saif Mohammad. 2012a. #emotional tweets. In **SEM 2012: The First Joint Conference on Lexical and Computational Semantics*.

Saif Mohammad. 2012b. 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*.

Olutobi Owoputi, Brendan OConnor, Chris Dyer, Kevin Gimpel, Nathan Schneider, and Noah A. Smith. 2013. Improved part-of-speech tagging for online conversational text with word clusters. In *Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL-2013)*.

W. Gerrod Parrott, editor. 2001. *Emotions in Social Psychology*. Psychology Press.

Matthew Purver and Stuart Battersby. 2012. Experimenting with distant supervision for emotion classification. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, EACL '12, pages 482–491.

Ashequl Qadir and Ellen Riloff. 2013. Bootstrapped learning of emotion hashtags #hashtags4you. In *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*.

Ellen Riloff, Ashequl Qadir, Prafulla Surve, Lalindra De Silva, Nathan Gilbert, and Ruihong Huang. 2013. Sarcasm as contrast between a positive sentiment and negative situation. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, EMNLP '13.

Kirk Roberts, Michael A. Roach, Joseph Johnson, Josh Guthrie, and Sanda M. Harabagiu. 2012. Empatweet: Annotating and detecting emotions on twitter. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012)*. ACL Anthology Identifier: L12-1059.

Jacopo Staiano and Marco Guerini. 2014. Depechemood: a lexicon for emotion analysis from crowd-annotated news. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*.

Carlo Strapparava and Rada Mihalcea. 2007. SemEval-2007 Task 14: Affective Text. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*.

Xiaolong Wang, Furu Wei, Xiaohua Liu, Ming Zhou, and Ming Zhang. 2011. Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, CIKM '11.

Wenbo Wang, Lu Chen, Krishnaprasad Thirunarayan, and Amit P. Sheth. 2012. Harnessing twitter "big data" for automatic emotion identification. In *Proceedings of the 2012 ASE/IEEE International Conference on Social Computing and 2012 ASE/IEEE*

International Conference on Privacy, Security, Risk and Trust, SOCIALCOM-PASSAT '12.

Changhua Yang, Kevin Hsin-Yih Lin, and Hsin-Hsi Chen. 2007a. Building emotion lexicon from weblog corpora. In *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*, ACL '07.

Changhua Yang, Kevin Hsin-Yih Lin, and Hsin-Hsi Chen. 2007b. Emotion classification using web blog corpora. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, WI '07, pages 275–278.