Steps Toward Automatic Understanding of the Function of Affective Language in Support Groups

Amit Navindgi* Veritas Technologies Mountain View, CA `navindgi@usc.edu`

Caroline Brun, Cécile Boulard Masson Xerox Research Centre Europe Meylan, France `{caroline.brun, cecile.boulard}@xrce.xerox.com`

Scott Nowson* Accenture Centre for Innovation Dublin, Ireland `scott.nowson@accenture.com`

Abstract

Understanding expression of emotions in support forums has great value and NLP methods are key to automating this. Many approaches use subjective categories which are more fine-grained than a straightforward polarity-based spectrum. However, the definition of such categories is non-trivial, and we argue for a need to incorporate communicative elements even beyond subjectivity. To support our position, we report experiments on a sentiment-labelled corpus of posts from a medical support forum. We argue that a more fine-grained approach to text analysis important, and also simultaneously recognising the social function behind affective expressions enables a more accurate and valuable level of understanding.

1 Introduction

There are a wealth of opinions on the internet. Social media has lowered the accessibility bar to an even larger audience who are now able to share their voice. However, more than just opinions on external matters, people are able to share their emotions and feelings, talking openly about very personal matters. Online presence has been shown to increase the chance of sharing personal information and emotions compared to face-to-face interactions (Hancock et al., 2007).

Medical support forums are one platform on which users generate emotion-rich content, exchange factual information about elements such as treatments or hospitals, and provide emotional support to others (Bringay et al., 2014). This sharing through open discussion is known to be considerably beneficial (Pennebaker et al., 2001).

Understanding affective language in the healthcare domain is an effective application of natural language technologies. Sentiment mining on platforms such as Twitter, for example, is a quick method to gauge public opinion of government policies (Speriosu et al., 2011). However, the level of affective expressions in a support forum setting is considerably more complex than a traditional positive-negative polarity spectrum.

More than just a more-fined grained labelling scheme, we also need a deeper understanding on the language being used. Much sentiment analysis research has focused on classifying the overall sentiment of documents onto a positive-negative spectrum (Hu and Liu, 2004). Recently, research work targeting finer grained analysis has emerged, such as aspect-based sentiment analysis (Liu, 2012; Pontiki et al., 2014), or semantic role labelling of emotions (Mohammad et al., 2014). This relatively new trend in social media analytics enables the detection of not simply binary sentiment, but more nuanced sentiments and mixed feelings. Such affective expressions often serve a social purpose (Rothman and Magee, 2016).

With this in mind, we explore a dataset drawn from a health-related support forum, labelled for a variety of expressed sentiments. Here, we do not necessarily seek state-of-the-art performance, but use this task to argue for two key positions:

- that sub-document level analysis is required to

*Work carried out while all authors at Xerox Research Centre Europe.

Proceedings of The Fourth International Workshop on Natural Language Processing for Social Media, pages 26–33, Austin, TX, November 1, 2016.

best understand affective expressions

- that to fully understand expressions of emotion in support forums, a fine-grained annotation scheme is required which takes into account the social function of such expressions.

This paper begins by reviewing work related to our propositions above. In Section 3 we describe the data which we have used, paying particular attention to the annotation scheme. We then report on our experiments, which were defined in order to support the hypotheses above. Following this, in Section 5 we discuss the implication of this work.

2 Related Work

As reported earlier, polarity-based studies in the healthcare domain have considerable value. One work squarely in the public policy domain sought to classify tweets related to the recent health care reform in the US into positive and negative (Speriosu et al., 2011). Ali et al. (2013) experimented with data from multiple forums for people with hearing-loss. They use the subjectivity lexicon of Wilson et al. (2005) and count-based syntactic features (e.g. number of adjectives, adverbs, etc.). This approach outperformed a baseline bag-of-words model, highlighting the importance of subjective lexica for text analysis in health domain. Ofek et al. (2013) use a dynamic sentiment lexicon to improve sentiment analysis in an online community for cancer survivors.

Sokolova and Bobicev (2013) took the lexicon approach further: they defined a more fine-grained annotation scheme (see Section 3) and labelled data from an IVF-related forum. Their category-specific set of lexicons performed better, at 6-class classification, than a generic subjectivity lexicon. In selecting their data, Sokolova and Bobicev (2013) – as Ali et al. (2013) and others have done – tapped into the domain of on-line support communities. Eastin and LaRose (2005) showed that people who seek support on-line – be it emotional or informational support – typically find it.

Informational support is based on sharing knowledge and experiences. Emotional support – framed as empathic communication – has four motivations: understanding, emotions, similarities and concerns (Pfeil and Zaphiris, 2007). In addition to direct support, another dimension of such online groups is self-disclosure (Prost, 2012). Barak and Gluck-Ofri (2007) identify self-disclosure as specific to open support groups (e.g. “Emotional Support for Adolescents”) as opposed to, for example, subject-specific discussion forums (e.g. “Vegetarianism and Naturalism” or “Harry Potter The Book”). Self-disclosure serves three social functions (Tichon and Shapiro, 2003): requesting implicit support by showing confusion and worries; providing support by sharing details of a personal experience and sharing information to further develop social relationships.

3 Data

3.1 Data Source

The data used here[1] is that of Bobicev and Sokolova (2015), an extension of the data described in Sokolova and Bobicev (2013). Data was collected from discussion threads on a sub-forum of an *In Vitro Fertilization (IVF)* medical forum[2] used by participants who belong to a specific age-group (over 35s). The dataset (henceforth *MedSenti*) originally contained 1321 posts across 80 different topics.

3.2 Annotation Details

There are two approaches to annotation of subjective aspects of communication: from the perspective of a reader’s perception (Strapparava and Mihalcea, 2007) or that of the author (Balahur and Steinberger, 2009). In labelling *MedSenti* Sokolova and Bobicev (2013) opted for the reader-centric model and hence asked the annotators to analyse a post’s sentiment as if they were other discussion participants. This is an important differentiation for automated classification style tasks - models are built to predict how people will understand the emotion expressed, as opposed to the emotion or sentiment an author feels they are conveying. The annotation scheme was evolved over multiple rounds of data exploration, and ultimately three sentiment categories were defined:

1. **confusion**, (henceforth `CONF`) which includes aspects such as “worry, concern, doubt, im-

[1] Kindly provided to us by the authors.

[2] http://ivf.ca/forums

patience, uncertainty, sadness, angriness, embarrassment, hopelessness, dissatisfaction, and dislike”

2. **encouragement**, (ENCO) includes “cheering, support, hope, happiness, enthusiasm, excitement, optimism”
3. **gratitude**, (GRAT) which represents thankfulness and appreciation

This set of labels captures important dimensions identified in the sociology literature. CONF here, for example, maps to expressions of confusion (Tichon and Shapiro, 2003) and those of concern (Pfeil and Zaphiris, 2007).

CONF is essentially a negative category while ENCO is positive. GRAT would therefore be a subset of positive expressions. In contrast, however, it was clear that certain expressions which might be considered negative on a word level – such as those of compassion, sorrow, and pity – were used with a positive, supportive intention. They were therefore included in the ENCO category, and were often posted with other phrases which would in isolation fall under this label.

In addition to the subjective categories, Sokolova and Bobicev (2013) identified two types of objective posts: those with strictly *factual* information (FACT), and those which combined factual information and short emotional expression (typically of the ENCO type) which were labelled as *endorsement* (ENDO). Each of the 1321 individual posts was labelled with one of the above five classes by two annotators.[3]

3.3 Data and Label Preprocessing

We select document labels as per Bobicev and Sokolova (2015): when two labels match, reduce to a single label; when the labels disagree the post is marked with a sixth label *ambiguous* (AMBI), which was not used in any experiment here. Posts with previous post quotation are annotated with (“QOTEHERE”), and quoting posts which contained no additional content were removed. This leaves 1137 posts in our *MedSenti* corpus, with the category distribution as per Table 1.

[3]Fleiss kappa = 0.73 (Bobicev and Sokolova, 2015).

Class	# Posts	%age	# Sents	%age
CONF	115	10.1	1087	13.5%
ENCO	309	27.2	1456	18.0%
ENDO	161	14.2	1538	19.1%
GRAT	122	10.7	733	9.1%
FACT	430	37.8	3257	40.4%
TOTAL	**1137**		**8071**	

Table 1: Class distribution of posts and sentences

4 Experiments

To support our positions for understanding affective expressions in support forums, and highlight some of the challenges with current approaches, we report a series of experiments.

4.1 Broad methodology

We use a robust dependency syntactic parser (Ait-Mokhtar et al., 2001) to extract a wide range of textual features, from n-grams to more sophisticated linguistic attributes. Our experiments are framed as multi-class classification tasks using liblinear (Fan et al., 2008) and used 5-fold stratified cross-validation. We do not use, here, a domain-tuned lexicon. We re-implemented the Health Affect Lexicon (Sokolova and Bobicev, 2013) and it performed as well as previously reported. However, such lexicons do not generalise well, and label-based tuning is very task specific. We use the current set of categories to make more general points about work in support-related domains.

4.2 Document Level analysis

Here, we consider each post as a single unit of text with a single label.

4.2.1 5-class classification

We utilised combinations of different linguistic feature sets, ranging from basic n-grams, through semantic dependency features. Here, we list the best performing combination: word uni-, bi-, and trigrams; binary markers for questions, conjunctions and uppercase characters; and a broad-coverage polarity lexicon. Results can be seen in Table 2

Our best overall score (macro averaged $F1 = 0.449$) is significantly above the majority class baseline ($F = 0.110$). This compares favorably with the six-class performance of semantic features of the original data analysis ($F1 = 0.397$, Sokolova and

	P	*R*	*F*
CONF	0.363	0.357	0.360
ENCO	0.555	0.854	0.673
ENDO	0.147	0.062	0.087
GRAT	0.583	0.492	0.533
FACT	0.573	0.502	0.535
MacroAvg	0.444	0.453	0.449

Table 2: Precision, Recall and F1 for the best feature set on 5-class document-level classification

Bobicev, 2013). However, more important – and not previously reported – is the per-category performance which gives more insight into the data. Essentially, we see that ENCO, GRAT and FACT perform relatively well while CONF and in particular ENDO are considerably poor.

To further explore this result we analyzed the error matrix (Navindgi et al., 2016). Looking at ENDO we see that incredibly only 6% has been correctly classified, while 86% is classified as either FACT or ENCO. This is theoretically understandable since the ENDO category is defined as containing aspects of both the other two categories directly. The reverse mis-classification is considerably less common, as is mis-classification as GRAT. CONF is also misclassified as FACT a majority, with 43%. One-vs-All analysis allows further insight (Navindgi et al., 2016). It is clear that this challenge is not a trivial one - there are distinct patterns of errors when classifying at the document level. In order to investigate this further, we move to sentence-level classification.

4.3 Sentence Level analysis

In sentence-level analysis, we tokenise each post into its constituent sentences. The 1137 *MedSenti* posts become 8071 sentences, *MedSenti-sent*. As manual annotation at sentence level would be too costly, we used automated methods to label the corpus with the five categories of sentiment.

4.3.1 Naïve Labelling

The most trivial approach to label sentences is for each sentence to inherit the label of the post in which it is present. Following this method, we obtain the distribution as reported in Table 1

We run the 5-class classification scenario on *MedSenti-sent* using the same conditions and the previous best feature set; the results are shown in Table 3. Overall, the performance is worse than the post-level counterpart, with the exception of a small improvement to ENDO. FACT is the best performing individual category, though now with greater recall than precision.

	P	*R*	*F*
CONF	0.235	0.157	0.188
ENCO	0.343	0.360	0.351
ENDO	0.174	0.088	0.117
GRAT	0.264	0.225	0.243
FACT	0.443	0.598	0.509
MacroAvg	0.291	0.286	0.289

Table 3: Precision, Recall and F1 for Sentence-level classification

We also explore the model performance with the error matrix (Navindgi et al., 2016). Our main observation is that the drop in performance of the four subjective categories is largely due to misclassification of sentences as FACT. Sentences in this category are the majority in *MedSenti-sent*. However, the proportional differences with *MedSenti* do not seem to be not enough to explain the significant changes.

A more likely explanation is simply that the errors arise because – at the very least – there can be FACT-like sentences in any post. At the time of creation, annotators were asked to label "the most dominant sentiment in the whole post" (Sokolova and Bobicev, 2013, p. 636). For example, post 141143 contains the sentence:

> Also, a nurse told me her cousin, 44, got pregnant (ivf)- the cousin lives in the USA.

The post itself is labelled ENCO. Strictly speaking, this sentence reports a fact, although it is easy to see how its purpose is to encourage others.

4.3.2 Subjectivity-informed labelling

One approach to re-labelling of data is to take advantage of coarser levels of annotation: that of subjectivity. Is it possible to at least distinguish which sentences are objective, and could be labelled as FACT? We have developed a subjectivity model[4] built for the SemEval 2016 Aspect Based Sentiment Analysis track (Pontiki et al., 2016), which

[4] brun-perez-roux:2016:SemEval

was among the top performing models for polarity detection. We ran this model on all sentences of the corpus in order to assess their subjectivity. Any sentence with a subjectivity likelihood of < 0.7 we consider to be *objective*; we also removed any *subjective* sentences which were previously FACT. This *MedSenti-sent-subj* set consists of 4147 sentences. We use the same experimental settings as previously, with results presented in Table 4.

	P	*R*	*F*
CONF	0.315	0.169	0.220
ENCO	0.390	0.457	0.421
ENDO	0.289	0.126	0.176
GRAT	0.284	0.294	0.289
FACT	0.543	0.745	0.628
MacroAvg	0.364	0.358	0.361

Table 4: Precision, Recall and F1 for Sentence-level classification of subjectivity-adjusted corpus

Performance is marginally better with this approach (against a majority macro averaged baseline of $F1 = 0.107$). Importantly, in analysing the error matrix[5] the proportion of data mis-classified has dropped considerably (from 51% to 37%). However, a related consequence is that the error-rate between the *subjective* categories has increased.

5 Discussion

Despite the disappointing results in our sentence level experiments, we maintain that this level of analysis, as a step toward aspect-based understanding, is important to explore further. One reason for poor performance with both the *MedSenti-sent* and *MedSenti-sent-subj* is the approach to annotation at the sentence level. Naturally manual annotation of 8K sentences is considerably expensive and time consuming. However, there are clear examples in the data set of distinct labels being required. Consider the following example, (with manually annotated, illustrative labels):

> post_id_226470 author1 "*author2 said [...]* <ENCO> Thanks,I think we were afraid of rushing into such a big decision but now I feel it is most important not to have regrets. </ENCO> <FACT> The yale biopsy is a biopsy of the lining of my uterus and it is a new test conducted by Yale University. Here is a link you can read: URL This test is optional and I have to pay for it on my own... no coverage.</FACT>"

[5]Not presented here for space concerns.

The first statement of this post is clearly intended to encourage the person to whom the author was responding. The second set of sentences is conveying deliberately factual information about their situation. In the *MedSenti* set this post is labelled as ENDO- the combination of ENCO and FACT. However, the FACT component of the post is a response to a question in an even earlier post than the quoted message. It could be argued therefore that these sentiment do not relate in the way for which the ENDO label was created. To consider post-level labels, then, we would argue is too coarse grained.

To explore the possible confusion introduced by the ENDO category, particularly after removing the objective sentences in *MedSenti-sent-subj*, we conducted experiments with this category. In this three-class experiment (ENCO, CONF, and GRAT), performance was again reasonable against baseline ($F1 = 0.510$ over $F1 = 0.213$), but the error rate was still high, particularly for GRAT. Regardless of the linguistic feature sets, the models do not appear to be capturing the differences between the subjective categories. This seems contradictory to the original authors' intention of building "a set of sentiments that [...] makes feasible the use of machine learning methods for automate sentiment detection." (Sokolova and Bobicev, 2013, p. 636). This is interesting because, from a human reader perspective (see Section 3), the annotation scheme makes intuitive sense. That the expressions of "negative" emotions such as sympathy be considered in the "positive" category of ENCO aligns with the social purpose behind such expressions (Pfeil and Zaphiris, 2007). Without explicitly calling attention to it, Sokolova and Bobicev (2013) encoded social purpose into their annotation scheme. As with previous effort in the space, the scheme they have defined is very much tuned to the emotional support domain.

In an attempt to understand potential reasons for errors, we created a visualisation of the annotation scheme in terms of scheme category label, higher level polarity, and sentiment target, which can be seen in Figure 1. As per the definitions of the cat-

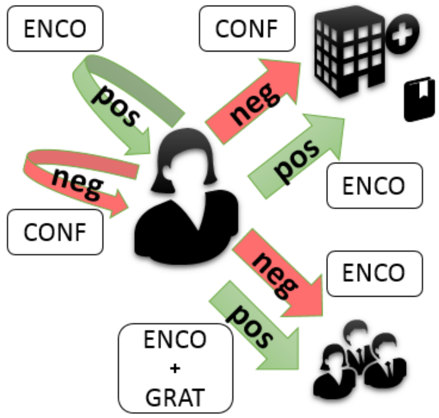

Figure 1: Visualisation of polarity to category mapping given affect target – one of either self, fellow forum participant, or external entity

egories, emotions expressed towards external entities, or oneself are clearly either positive-`ENCO` or negative-`CONF`. However, the pattern is different in interpersonal expression between forum contributors. In the medical support environment "negative" expressions, as previously discussed serve a positive and supportive purpose. Also, the category of `GRAT`– a positive expression – is always in this situation directed to another participant. This makes the interpersonal expression loadings both overloaded both in terms of classification and polarity. These relationships, in many ways, make machine modelling therein overly noisy.

Of course, it is fair to say that one direction of work in such a social domain that we did not explore is context. The original authors report subsequently on incorporating context into their experiments: both in terms of the position within a discussion of a post (Bobicev and Sokolova, 2015) and the posting history of an author (Sokolova and Bobicev, 2015). In this work we have eschewed context, though acknowledge that it is significantly important: in the `ENCO-FACT` sample above, for example, context may enable a better understanding that the `ENCO` sentence is in response to another `ENCO` statement, while the `FACT` is a response to a direct question. In this sense, there is a clear motivation to understand document-level relationships at the sentence level.

Another direction which could be explored is an alternative annotation scheme. Prost (2012) suggests an annotation scheme used to identify the sharing of both practice-based and emotional support among participants of online forums for teachers. This annotation scheme is a combination of schemes developed for social support forums with those created for practice-based forums. Identifed categories and sub-categories are described in Table 5.

Category	Subcategory
Self disclosure	professional experience personal experience emotional expression support request
Knowledge sharing	from personal experience Concrete info or documents
Opinion/evaluation	*na*
Giving advice	*na*
Giving emotional support	*na*
Requesting clarification	*na*
Community building	reference to community humour broad appreciation direct thanks
Personal attacks	*na*

Table 5: Categories and subcategories from support annotation scheme of Prost (2012)

Most of the categories are relevant for both types of forums, support and practice-based. Prost annotated texts at the sub-sentence level, with these 15 categories. In order to produce the volumes of data that would be necessary for machine-learning based approaches to understanding support forum, this is impractical. There is clearly a balance to be struck between utility and practicality. However, Prost's scheme illustrates that in sociological circles, it is important to consider the social context of subjective expressions: there are two categories equivalent to `GRAT` here, one which is more directed, and the other which concerns a bigger picture expression of the value of community.

6 Conclusion

In this work we have argued two positions. Despite seemingly poor results at sentence-level, we are convinced that the examples we have provided demonstrate that document-level analysis is insufficient to accurately capture expressions of sentiment in emotional support forums. We have also shown that there are important social dimensions to this type of domain which should also be taken into account. It is clear that there is considerable value to be gained from automated understanding of this increasing body of data; we in the Social NLP community need to consider some more refined approaches in order to maximise both the value itself and its fidelity.

References

Salah Ait-Mokhtar, Jean-Pierre Chanod, and Claude Roux. 2001. A multi-input dependency parser. In *Proceedings of the Seventh International Workshop on Parsing Technologies*.

Tanveer Ali, David Schramm, Marina Sokolova, and Diana Inkpen. 2013. Can i hear you? sentiment analysis on medical forums. In *Proceedings of the sixth international joint conference on natural language processing, Asian Federation of Natural Language Processing, Nagoya, Japan, October 2013*, pages 667–673.

Alexandra Balahur and Ralf Steinberger. 2009. Rethinking sentiment analysis in the news: from theory to practice and back. In *Proceedings of the 1st Workshop on Opinion Mining and Sentiment Analysis, 2009*.

Azy Barak and Orit Gluck-Ofri. 2007. Degree and reciprocity of self-disclosure in online forums. *Cyberpsychology & Behavior*, 10(3):407–417.

Victoria Bobicev and Marina Sokolova. 2015. No sentiment is an island - sentiment classification on medical forums. In Nathalie Japkowicz and Stan Matwin, editors, *Discovery Science - 18th International Conference, DS 2015, Banff, AB, Canada, October 4-6, 2015, Proceedings*, volume 9356 of *Lecture Notes in Computer Science*, pages 25–32. Springer.

Sandra Bringay, Eric Kergosien, Pierre Pompidor, and Pascal Poncelet, 2014. *Identifying the Targets of the Emotions Expressed in Health Forums*, pages 85–97. Springer Berlin Heidelberg, Berlin, Heidelberg.

Matthew S. Eastin and Robert LaRose. 2005. Alt.support: modeling social support online. *Computers in Human Behaviour*, 21(6):977–992.

Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. 2008. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9:1871–1874.

Jeffrey T. Hancock, Catalina Toma, and Nicole Ellison. 2007. The truth about lying in online dating profiles. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 449–452, New York, NY, USA. ACM.

Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *KDD*, pages 168–177.

Bing Liu. 2012. *Sentiment Analysis and Opinion Mining*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers.

Saif Mohammad, Xiaodan Zhu, and Joel Martin, 2014. *Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, chapter Semantic Role Labeling of Emotions in Tweets, pages 32–41. Association for Computational Linguistics.

Amit Navindgi, Caroline Brun, Cécile Boulard Masson, and Scott Nowson. 2016. Toward automatic understanding of the function of affective language in support groups. *Computation and Language, arXiv preprint, http://arxiv.org/abs/1610.01910*.

Nir Ofek, Cornelia Caragea, Lior Rokach, and Greta E Greer. 2013. Improving sentiment analysis in an online cancer survivor community using dynamic sentiment lexicon. In *Social Intelligence and Technology (SOCIETY), 2013 International Conference on*, pages 109 – 113.

James W. Pennebaker, Emmanuelle Zech, and Bernard Rim, 2001. *Disclosing and sharing emotion: Psychological, social, and health consequences*, pages 517–539. American Psychological Association.

Ulrike Pfeil and Panayiotis Zaphiris. 2007. Patterns of empathy in online communication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI07)*, pages 919–928, San Jose, CA, USA.

Maria Pontiki, Dimitrios Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. Semeval-2014 task 4: Aspect based sentiment analysis. In *International Workshop on Semantic Evaluation (SemEval)*.

Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammad AL-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphee De Clercq, Veronique Hoste, Marianna Apidianaki, Xavier Tannier, Natalia Loukachevitch, Evgeniy Kotelnikov, Núria Bel, Salud María Jiménez-Zafra, and Gülşen Eryiğit. 2016. Semeval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 19–30, San Diego, California, June. Association for Computational Linguistics.

Magali Prost. 2012. *changes entre professionnels de l'ducation sur les forums de discussion: entre soutien psychologique et acquisition de connaissances sur la pratique*. Ph.D. thesis, Telecom ParisTech, Paris, France.

Naomi B. Rothman and Joe C. Magee. 2016. Affective expressions in groups and inferences about members' relational well-being: The effects of socially engaging and disengaging emotions. *Cognition & Emotion, Special Issue on Emotions in Groups*, 30(1):150–166.

Marina Sokolova and Victoria Bobicev. 2013. What sentiments can be found in medical forums? In Galia Angelova, Kalina Bontcheva, and Ruslan Mitkov, editors, *Recent Advances in Natural Language Processing, RANLP 2013, 9-11 September, 2013, Hissar, Bulgaria*, pages 633–639. RANLP 2013 Organising Committee / ACL.

Marina Sokolova and Victoria Bobicev. 2015. Learning relationship between authors' activity and sentiments: A case study of online medical forums. In Galia Angelova, Kalina Bontcheva, and Ruslan Mitkov, editors, *Recent Advances in Natural Language Processing, RANLP 2015, 7-9 September, 2015, Hissar, Bulgaria*, pages 604–610. RANLP 2015 Organising Committee / ACL.

Michael Speriosu, Nikita Sudan, Sid Upadhyay, and Jason Baldridge. 2011. Twitter polarity classification with label propagation over lexical links and the follower graph. In *Proceedings of the First workshop on Unsupervised Learning in NLP*, EMNLP '11, pages 53–63.

Carlo Strapparava and Rada Mihalcea. 2007. Semeval-2007 task 14: Affective text. In *Proceedings of the 2008 ACM symposium on Applied computing, 2008*.

Jennifer G. Tichon and Margaret Shapiro. 2003. The process of sharing social support in cyberspace. *Cyberpsychology & Behavior*, 6(2):161–170.

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proc. of HLT-EMNLP-2005*.