Fine-Grained Emotion Detection in Health-Related Online Posts

Hamed Khanpour
Computer Science and Engineering
University of North Texas
Denton, TX 76207
`hamed.khanpour@microsoft.com`

Cornelia Caragea
Computer Science
Kansas State University
Manhattan, KS 66502
`ccaragea@ksu.edu`

Abstract

Detecting fine-grained emotions in online health communities provides insightful information about patients' emotional states. However, current computational approaches to emotion detection from health-related posts focus only on identifying messages that contain emotions, with no emphasis on the emotion type, using a set of handcrafted features. In this paper, we take a step further and propose to detect fine-grained emotion types from health-related posts and show how high-level and abstract features derived from deep neural networks combined with lexicon-based features can be employed to detect emotions.

1 Introduction

Emotions are an essential part of our lives and reflect feelings such as joy, sadness, and fear, which affect our overall wellbeing. Emotion detection from text using computational models has been extensively studied from data such as news headlines, social media, blog posts, and song lyrics (Katz et al., 2007; Abdul-Mageed and Ungar, 2017; Mohammad and Turney, 2013; Strapparava and Mihalcea, 2007; Liew and Turtle, 2016).

Recently, emotion detection started to emerge in online health communities (OHCs) (Khanpour et al., 2018; Biyani et al., 2014; Wang et al., 2012). OHCs provide a user-friendly environment for patients, their families and friends to share thoughts and socialize with each other on topics such as therapeutic processes, side effects, and mental and emotional health. Emotion detection in OHCs is substantially different from the general text due to a health-related vocabulary that people use in OHCs. For example, the phrase "hot flashes" may not have a specific meaning in the general domain, but it bears a very negative feeling in patients' posts. Similarly, the post: "Just received

"My doctor's office is very clean, who cares when he has prescribed me a wrong medication for six months!"

Table 1: Example of an emotional message from an OHC that contains a sad emotion.

notice from my doctor, I have a positive Cologuard result" is associated with a very sad emotion in a health domain, i.e., when a test is positive, it means that the disease is present.

Despite of emergence of emotion detection in OHCs, most of the recent works have been devoted to high level emotion analysis, i.e., identifying messages that contain emotions, with no emphasis on the unique challenges associated with fine-grained emotion detection. In order to correctly detect the types of emotions present in messages posted in OHCs, a deep understanding of the text and the writer's intention are required. Table 1 shows an example of a message that contains a sad emotion, which is hidden in text. We ran several sentiment tools on this message, including Stanford CoreNLP (Socher et al., 2013), and interestingly, all showed a positive sentiment, while the emotion expressed is clearly one of sadness (mixed with sarcasm). Thus, even an approach to predict the negative sentiment of the message would not suffice.

In this paper, we analyze messages in OHCs to understand the most prominent emotions in health-related posts and propose a computational model that is able to exploit the semantic information from text and coherently combines high-level (abstract) features with surface and lexicon-based features to automatically detect fine-grained emotions. Our contributions are as follows:

1. We study fine-grained emotions and their distribution in messages posted in OHCs by constructing and analyzing two health-related datasets for fine-grained emotion detection.

Identifying emotions in patients' messages augments the capability of OHCs' moderators, caregivers, and doctors to provide high-quality services to OHCs' users and patients. To our knowledge, we are the first to address fine-grained emotion detection in OHCs.

2. We propose a computational model, called ConvLexLSTM, for emotion detection in OHCs. Our model combines the output of a Convolutional Neural Network (CNN) with lexicon-based features, which are all fed into a Long Short-Term Memory (LSTM) network that produces the final output via softmax. We show empirically that ConvLexLSTM significantly outperforms strong baselines and prior works. Moreover, we show that the proposed model continues to perform well even in the absence of lexicon features.

3. Finally, we apply ConvLexLSTM in a large scale experiment to study the correlation between US holidays and users' emotional states, which can help design smarter approaches to improve patients' moods.

2 Related Work

Emotion detection has been extensively studied in computational linguistics for a long time (Mohammad and Turney, 2013; Strapparava et al., 2004). The Ekman's basic emotion set, which includes six emotions: anger, disgust, fear, happiness, sadness, and surprise, is arguably the best well-known emotion categorization (Ekman, 1992). Strapparava and Mihalcea (2008) proposed knowledge-based and corpus-based methods for classifying emotions based on Ekman's basic emotions. Co-occurrence of general words with emotional words has beed used by Katz et al. (2007) for identifying emotion types latent in news headlines. Keyword-based approaches that are based on finding emotional words in text suffer from the inability to classify text that lacks specific keywords. Therefore, Bao et al. (2009) proposed to use topical relations between words and emotion types for emotion classification in online news. Strapparava et al. (2012) studied emotions from song lyrics.

Emotion detection has been studied in social media as well and brings additional challenges due to their informal context in which people do not follow grammatical rules and use many characters that do not occur in formal text (e.g., #, :)). Emotion lexicons derived from social media,

e.g., based on *emotion word hashtags*, have been shown to improve models' performance for emotion detection (Mohammad, 2012; Sykora et al., 2013). Abdul-Mageed and Ungar (2017) used distant supervision to construct a large dataset from the general Twitter for fine-grained emotion detection and explored deep learning models to detect emotions. Liew and Turtle (2016) and Liew et al. (2016) also created a dataset of about $15,500$ tweets labeled with 28 emotion types, using the Amazon Mechanical Turk. Some studies combined knowledge-based and keyword-based approaches with linguistic features and used a machine learning algorithm to reasonably classify sentences with no emotional keywords (Yang et al., 2012; Neviarouskaya et al., 2010).

Emotional support is considered as one of the main advantages of using OHCs that brings better feelings (Khanpour et al., 2017; Zhang et al., 2017a,b; Qiu et al., 2011) and fewer mortality odds to patients (Kroenke et al., 2013). Interestingly, to this date, only very few studies have started to analyze emotions in OHCs using computational models (Wang et al., 2012; Biyani et al., 2014; Wang et al., 2014b,a). For example, Wang et al. (2012) used linear regression to identify emotional support from a cancer forum. Their model predicts to what extent each sentence contains either emotional or non-emotional support. Their features include: LIWC features, POS tags, message length, subjectivity intensity, and LDA topical features. Biyani et al. (2014) identified messages that contain emotions in a breast cancer forum using unigrams, POS tags, structural patterns, and five lexicons that contain strong and weak subjective words, cancer drugs, side-effects, and cancer procedures, and showed that lexicons features have a high impact on the results. Khanpour and Caragea (2018) used deep learning to extract therapeutic processes and side effects from patients' posts. Wang et al. (2014b) classified OHCs' messages based on the intention of the participant when writing messages (e.g., seeking or providing information) and used a combination of features from Wang et al. (2012) coupled with lexicon features used in Biyani et al. (2014).

3 Data Collection and Annotation

To study the most prominent emotions and their distribution in OHCs and to evaluate our model for fine-grained emotion detection in OHCs, we

constructed two benchmark datasets, since, to our knowledge, no labeled dataset is available for this task. The first dataset is created by using data from Biyani et al. (2014), which contains 1066 sentences from the breast cancer discussion board in the Cancer Survivors' Network (CSN) of the American Cancer Society, denoted as B-DS. Note that Biyani et al. (2014) performed sentence level classification since longer messages often comprise different topics. A sentence level analysis can make a better estimation on the purpose of the commentator in writing that sentence, whether or not expressing his or her emotions. For the second dataset, we randomly selected 225 comments from 21 discussion threads in the lung cancer discussion board of CSN. We denote this second dataset as L-DS. Following Biyani et al. (2014), we extracted all sentences out of comments and chose sentences with a length greater than four words. We ended up with 1041 sentences in L-DS. In total, we have 2107 sentences labeled with emotion types.

For our annotation task, we followed the six emotions suggested by Ekman (1992). Our annotators were allowed to attribute one or more emotions to a single sentence, e.g., a sentence could be annotated as bearing *sadness* or a combination of *sadness* and *fear*. The annotation task was conducted iteratively following prior studies and guidelines, using three training rounds (DMello, 2016; Fort et al., 2016; Shanahan et al., 2006). In each round, 300 sentences drawn from both B-DS and L-DS were assigned to the annotators,[1] and we asked them to meet with the researchers in a group to discuss disagreements and document their discussion before the next 300 instances were assigned. Upon passing the training period, annotators were assigned to annotate the remaining sentences from B-DS and L-DS, and they ended up with 83% Kappa inter-annotator agreement. For the remaining 17%, the annotators expressed their views on each case in the presence of the researchers and finally 100% agreement was achieved during 20 days. All these sentences plus the 900 ones used during the three training rounds became part of the final dataset. Table 2 represents the distribution of emotions in 2107 sentences. Note that some sentences do not contain any emotion. As can be seen from the table, both datasets have a similar distribution of emotions, with joy and sadness being the most prominent.

---

[1]Annotators were three graduate students.

Emotions	Lung	Breast	Percentage(%)
Anger	59	29	4.0
Disgust	4	1	0.2
Fear	33	39	3.4
Joy	368	470	39.7
Sadness	183	134	15.0
Surprise	8	3	0.5

Table 2: Emotion distributions in B-DS and L-DS with the percentage estimated from both L-DS and B-DS.

4 Model

Given a sentence of n words, we apply CNN to extract high-level (abstract) features that capture the semantic part of the text (Lai et al., 2015). We combine high-level features with surface-level and lexicon-based features. Our proposed model, ConvLexLSTM, is shown graphically in Figure 1. As can be seen from the figure, we use a combination of CNN and LSTM models, where the final feature vectors from CNN augmented with lexicon-based features are fed as input to the LSTM network. The architecture of our proposed classification model is close to the models described in Kim et al. (2016) and Xiao and Cho (2016), in which they applied a character-level CNN to create high-level features, whereas our model works at word-level and uses lexicon features. We use the word-level input to benefit from applying embedding vectors, trained on OHCs. We use the character-level model by Kim et al. (2016), denoted by C-ConvLSTM, as one of our baselines.
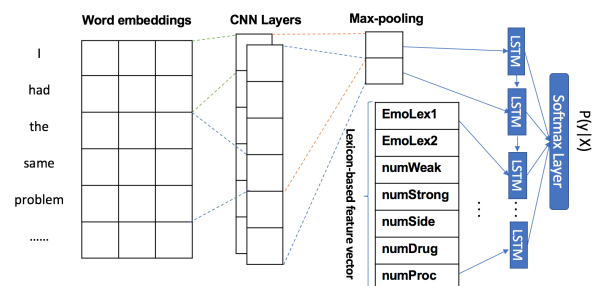

Figure 1: The structure of ConvLexLSTM.

Lexicon-based Features: Lexicon-based approaches for detecting emotions in the text have been the main stream of many models (Strapparava et al., 2004; Strapparava and Mihalcea, 2008; Mohammad, 2012; Liu, 2012).

In our work, we used the same lexicons that were provided by Biyani et al. (2014) such as weak subjective words (**numWeak**), strong subjective words (**numStrong**) (Stoyanov et al., 2005), cancer drugs (**numDrug**), side-effects (**numSide**), and therapeutic procedures (**numProc**). These

lexicons address differentiation between emotional versus non-emotional messages. However, we need more granular information for differentiating between a variety of emotion types. Hence, we also used lexicons introduced by Strapparava and Mihalcea (2007), denoted as **EmoLex1**, and by Mohammad and Turney (2013), denoted as **EmoLex2**. We use frequencies of lexicon words to construct the lexicon-based feature vectors. Note that prior work (Biyani et al., 2014) showed that LIWC did not generate high quality features for classification in OHCs, and thus, we did not use it in our study.

5 Experiments and Results

Next, we describe the evaluation of ConvLexLSTM, using the *joy* and *sadness* emotions, which have at least 5% coverage in our labeled data (see Table 2). Also, since binary tasks are considered easier to learn than multi-class tasks (Bishop, 2006), we trained our models in the two-class setting: joy/non-joy (and sad/non-sad), by binarizing the datasets. In all experiments, we used word embeddings as input to the neural networks, which were generated with the W2vector module in Gensim (Řehůřek and Sojka, 2010) on the data (users' comments) from all discussion boards of the Cancer Survivors' Network (CSN) of the American Cancer Society, between June 2000 and June 2012. We also experimented with word embeddings generated using Wikipedia, but these embeddings resulted in lower performance as compared with those generated using CSN data. We estimated hyper-parameters for each deep neural network via a grid search over combinations of important hyper-parameters (e.g., learning rate, decay rate, dropout, number of layers, filter region size, and number of filters). The grid search was done on a development set that consists of removing 20% of instances from the training set in each iteration of 10-fold cross-validation. We report precision, recall and F1-score.

5.1 Performance of ConvLexLSTM

First, we evaluate ConvLexLSTM performance in an ablation experiment to determine the role played by each component for emotion detection. Specifically, we compare ConvLexLSTM with ConvLSTM (a model that has the same architecture as ConvLexLSTM, but does not use any external lexicon), CNN, LSTM, and support vec-

	Method	B-DS			L-DS		
		Pr	Re	F1	Pr	Re	F1
Joy	ConvLexLSTM	**92.3**	**94.3**	**93.2**	**90.4**	**89.3**	**89.8**
	ConvLSTM	86.6	88.4	87.4	87.0	83.0	85.0
	CNN	85.0	84.0	84.5	82.2	82.8	82.5
	LSTM	86.0	86.6	86.3	85.0	83.0	84.0
	Seven-Lexicon	63.4	87.3	73.45	60.0	85.1	70.37
	C-ConvLSTM	86.2	87.0	86.6	85.0	82.0	83.47
	SWAT	66.0	68.0	67.0	65.5	66.7	66.0
	EmoSVM	81.0	82.0	81.5	82.0	80.0	81.0
Sad	ConvLexLSTM	**93.7**	**91.1**	**92.3**	**88.0**	**90.9**	**89.4**
	ConvLSTM	89.0	87.8	88.4	81.0	87.5	84.0
	CNN	83.2	83.6	83.4	81.7	80.5	81.0
	LSTM	87.4	85.8	86.6	83.2	83.6	83.4
	Seven-Lexicon	61.0	84.9	70.99	61.0	83.3	70.42
	C-ConvLSTM	85.0	83.6	84.3	83.7	82.1	82.9
	SWAT	65.0	66.0	65.5	64.0	65.0	64.5
	EmoSVM	80.5	81.7	81.0	79.0	78.0	78.5

Table 3: Emotion detection results using 10-fold cross validation. The numbers are percentages.

tor machine (SVM) with the (concatenated) features from the seven lexicons (described above).

Table 3 shows the results of this comparison. As can be seen from the table, ConvLexLSTM achieves the best results consistently throughout all experiments in terms of all compared measures. This ablation experiment confirms our intuition that all components are contributing to the final emotion detection. For example, removing the seven lexicon features from ConvLexLSTM, which yields ConvLSTM, results in a drop in F1-score by 5.8% on *joy* in B-DS, and by 3.9% on *sadness* in B-DS. Still, ConvLSTM is the second performing model in terms of F1-score. These results show that our model can be successfully applied in a health domain even in the absence of health lexicons, which are often expensive to obtain. Not surprisingly, the SVM with the seven-lexicon based features (denoted as Seven-Lexicon) performs the worst among the compared models, suggesting that capturing the semantic information from text via deep neural networks improves emotion detection.

Second, we compare ConvLexLSTM with three baselines: C-ConvLSTM (i.e., a character-level CNN-LSTM) (Kim et al., 2016), SWAT (Katz et al., 2007) (i.e., an emotion detection model from SemEval-2007), and EmoSVM (i.e., an SVM with a set of handcrafted features: unigrams, bigrams, POS tags, the word-emotions association lexicon by Mohammad (2012), the WordNet-Affect lexicon by Strapparava et al. (2004), and the output of the Stanford sentiment tool by Socher et al. (2013). Table 3 shows the results of this comparison as well. As can be seen, ConvLexLSTM out-
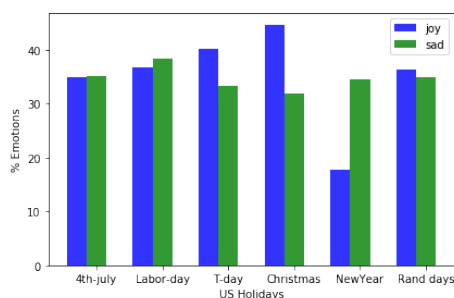
Figure 2: Joy and sadness throughout events.

performs all three baselines on both datasets, and more importantly, the character-level CNN-LSTM by Kim et al. (2016) (i.e., the C-ConvLSTM model). This result confirms our belief that applying word embedding vectors, which are trained directly on data from OHCs yields improvement in performance over character-level models.

It is worth mentioning that all deep neural networks, ConvLexLSTM, ConvLSTM, CNN, LSTM, and C-ConvLSTM, that capture high-level semantic features perform better than the traditional models on emotion detection. The lexicon-based features act as a complement (for the high-level semantic features) by looking into exact words in the text to generate appropriate features in ConvLexLSTM for emotion detection. With a paired T-test, the improvements of ConvLexLSTM over the compared models for F1-score are statistically significant for p-values < 0.05.

5.2   Impact of holidays on emotional states

We further analyzed the impact of several US holidays, i.e., 4th of July, Labor Day, Thanksgiving Day, Christmas Day, and New Year's Eve, on CSN users' emotional states, joy and sadness. For this experiment, we extracted messages from each event day itself and from two days before and two days after each event (holiday) from the entire CSN data. We also collected messages written on five random days (June 15, October 20, March 9, April 3, and January 29), which are not close to any event, to be used as a baseline for comparing the emotional states of participants in different holidays. Consistent with our labeled datasets, we removed messages with less than four tokens. We used our best performing model ConvLexLSTM to classify joy and sadness on these data.

Figure 2 shows the percentage of joy and sad predicted messages for each holiday and for the five random days. As can be seen from the figure, on the random days, the percentage of joy and sad emotions are similar, and so are they for 4th of July and Labor Day, whereas Christmas and Thanksgiving show more joyful spirits, possibly due to family gatherings and other social events around these holidays, in which people feel supported, and hence, feel better. Christmas shows an increase in joy and a slight decrease in sad emotions compared with Thanksgiving. Interestingly, around the New Year's Eve, the percentage of sad messages is almost twice higher compared with the percentage of joy messages, which can be attributed to the end of the holiday season and family gatherings and the beginning of a new challenging year.

6   Conclusion

In this paper, we addressed the problem of fine-grained emotion detection from OHCs messages. To this end, we first annotated a dataset from a cancer forum (i.e., the Cancer Survivors' Network of the American Cancer Society) with the six most common emotions suggested by Ekman (1992) and studied the most prominent emotions and their distribution in OHCs. We found that joy and sadness occur most frequently in the forum, followed by anger and fear. Not surprisingly, disgust and surprise appear the least number of times. We then proposed a computational model that combines the strengths of CNNs, LSTMs and lexicon-based approaches to capture the hidden semantics in OHCs messages and to provide a more insightful understanding of emotional messages by identifying their emotion types.

Our results are promising and show that our proposed model, with or without lexicon-based features, which are often expensive to obtain or maintain in a health domain, provides a better emotion type detection compared with strong baselines and prior works. Given our initial success, in the future, it would be valuable to construct a large health-related dataset to cover other types of emotions, e.g., anger or fear.

Acknowledgments

References

Muhammad Abdul-Mageed and Lyle Ungar. 2017. Emonet: Fine-grained emotion detection with gated recurrent neural networks. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 718–728. Association for Computational Linguistics.

Shenghua Bao, Shengliang Xu, Li Zhang, Rong Yan, Zhong Su, Dingyi Han, and Yong Yu. 2009. Joint emotion-topic modeling for social affective text mining. In *Data Mining, 2009. ICDM'09.*, pages 699–704.

Christopher M. Bishop. 2006. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc.

Prakhar Biyani, Cornelia Caragea, Prasenjit Mitra, and John Yen. 2014. Identifying emotional and informational support in online health communities. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 827–836, Dublin, Ireland. Dublin City University and Association for Computational Linguistics.

Sidney K DMello. 2016. On the influence of an iterative affect annotation approach on inter-observer and self-observer reliability. *IEEE Transactions on Affective Computing*, 7(2):136–149.

Paul Ekman. 1992. An argument for basic emotions. *Cognition & emotion*, 6(3-4):169–200.

Karën Fort et al. 2016. *Collaborative Annotation for Reliable Natural Language Processing: Technical and Sociological Aspects*. Wiley Online Library.

Phil Katz, Matthew Singleton, and Richard Wicentowski. 2007. Swat-mp: the semeval-2007 systems for task 5 and task 14. In *4th international workshop on semantic evaluations*, pages 308–313.

Hamed Khanpour and Cornelia Caragea. 2018. Fine-grained information identification in health related posts. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, SIGIR '18, pages 1001–1004.

Hamed Khanpour, Cornelia Caragea, and Prakhar Biyani. 2017. Identifying empathetic messages in online health communities. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing, IJCNLP 2017, Taipei, Taiwan, November 27 - December 1, 2017, Volume 2: Short Papers*, pages 246–251.

Hamed Khanpour, Cornelia Caragea, and Prakhar Biyani. 2018. Identifying emotional support in online health communities. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, New Orleans, Louisiana, USA, February 2-7, 2018*.

Yoon Kim, Yacine Jernite, David Sontag, and Alexander M. Rush. 2016. Character-aware neural language models. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, AAAI'16, pages 2741–2749. AAAI Press.

Candyce H Kroenke, Charles Quesenberry, Marilyn L Kwan, Carol Sweeney, Adrienne Castillo, and Bette J Caan. 2013. Social networks, social support, and burden in relationships, and mortality after breast cancer diagnosis in the life after breast cancer epidemiology (lace) study. *Breast cancer research and treatment*, 137(1):261–271.

Siwei Lai, Liheng Xu, Kang Liu, and Jun Zhao. 2015. Recurrent convolutional neural networks for text classification. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, AAAI'15, pages 2267–2273. AAAI Press.

Jasy Suet Yan Liew and Howard R. Turtle. 2016. Exploring fine-grained emotion detection in tweets. In *Proceedings of the Student Research Workshop, SRW@HLT-NAACL 2016, The 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego California, USA, June 12-17, 2016*, pages 73–80.

Jasy Suet Yan Liew, Howard R. Turtle, and Elizabeth D. Liddy. 2016. Emotweet-28: A fine-grained emotion corpus for sentiment analysis. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Paris, France. European Language Resources Association (ELRA).

Bing Liu. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1):1–167.

Saif M Mohammad. 2012. # emotional tweets. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*, pages 246–255. Association for Computational Linguistics.

Saif M Mohammad and Peter D Turney. 2013. Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, 29(3):436–465.

Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka. 2010. Emoheart: conveying emotions in second life based on affect sensing from text. *Advances in Human-Computer Interaction*, 2010:1.

Baojun Qiu, Kang Zhao, Prasenjit Mitra, Dinghao Wu, Cornelia Caragea, John Yen, Greta E Greer, and Kenneth Portier. 2011. Get online support, feel better–sentiment analysis and dynamics in an online cancer survivor community. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, pages 274–281. IEEE.

Radim Řehůřek and Petr Sojka. 2010. Software Framework for Topic Modelling with Large Corpora. In *LREC*, pages 45–50, Valletta, Malta.

James G Shanahan, Yan Qu, and Janyce Wiebe. 2006. *Computing attitude and affect in text: Theory and applications*, volume 20. Springer.

Richard Socher, Alex Perelygin, Jean Wu, Jason Chuang, Christopher D. Manning, Andrew Ng, and Christopher Potts. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1631–1642, Seattle, Washington, USA. Association for Computational Linguistics.

Veselin Stoyanov, Claire Cardie, and Janyce Wiebe. 2005. Multi-perspective question answering using the opqa corpus. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 923–930. Association for Computational Linguistics.

Carlo Strapparava and Rada Mihalcea. 2007. Semeval-2007 task 14: Affective text. In *4th Semantic Evaluations*, pages 70–74.

Carlo Strapparava and Rada Mihalcea. 2008. Learning to identify emotions in text. In *ACM Applied computing*, pages 1556–1560.

Carlo Strapparava, Rada Mihalcea, and Alberto Battocchi. 2012. A parallel corpus of music and lyrics annotated with emotions. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation, LREC 2012, Istanbul, Turkey, May 23-25, 2012*, pages 2343–2346.

Carlo Strapparava, Alessandro Valitutti, et al. 2004. Wordnet affect: an affective extension of wordnet. In *LREC*, volume 4, pages 1083–1086.

Martin D Sykora, Thomas Jackson, Ann O'Brien, and Suzanne Elayan. 2013. Emotive ontology: Extracting fine-grained emotions from terse, informal messages.

Xi Wang, Kang Zhao, and Nick Street. 2014a. Predicting user engagement in online health communities based on social support activities. In *Ninth INFORMS Workshop on Data Mining and Analytics*.

Xi Wang, Kang Zhao, and Nick Street. 2014b. Social support and user engagement in online health communities. In *Smart Health*, pages 97–110. Springer.

Yi-Chia Wang, Robert Kraut, and John M Levine. 2012. To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups. In *CSCW*, pages 833–842. ACM.

Yijun Xiao and Kyunghyun Cho. 2016. Efficient character-level document classification by combining convolution and recurrent layers. *arXiv preprint arXiv:1602.00367*.

Hui Yang, Alistair Willis, Anne De Roeck, and Bashar Nuseibeh. 2012. A hybrid model for automatic emotion recognition in suicide notes. *Biomedical informatics insights*, 5(Suppl 1):17.

Shaodian Zhang, Tian Kang, Lin Qiu, Weinan Zhang, Yong Yu, and Noémie Elhadad. 2017a. Cataloguing treatments discussed and used in online autism communities. In *Proceedings of the 26th International Conference on World Wide Web*, pages 123–131. International World Wide Web Conferences Steering Committee.

Shaodian Zhang, Erin OCarroll Bantum, Jason Owen, Suzanne Bakken, and Noémie Elhadad. 2017b. Online cancer communities as informatics intervention for social support: conceptualization, characterization, and impact. *Journal of the American Medical Informatics Association*, 24(2):451–459.