# Identifying Empathetic Messages in Online Health Communities

**Hamed Khanpour**
Computer Science and Engineering
University of North Texas
Hamedkhanpour@my.unt.edu

**Cornelia Caragea**
Computer Science
Kansas State University
ccaragea@k-state.edu

**Prakhar Biyani**
Yahoo Research
pxb5080@yahoo-inc.com

## Abstract

Empathy captures one's ability to correlate with and understand others' emotional states and experiences. Messages with empathetic content are considered as one of the main advantages for joining online health communities due to their potential to improve people's moods. Unfortunately, to this date, no computational studies exist that automatically identify empathetic messages in online health communities. We propose a combination of Convolutional Neural Networks (CNN) and Long Short Term Memory (LSTM) networks, and show that the proposed model outperforms each individual model (CNN and LSTM) as well as several baselines.

## 1 Introduction

Empathy captures the ability of an individual to correlate with and gain an accurate understanding of other individuals' emotional states by putting oneself in their situations with appropriate reactions (Batson, 2009; Launay et al., 2015). Empathy is shown to have a fundamental role in connecting people in a community together (Davis et al., 2004). Recently, many studies in social and psychological sciences have investigated the correlation between the empathetic capability of users in a social network and their characteristics and behavioral patterns. For example, Kardos et al. (2017) analyzed social networks and found that higher empathetic abilities in social network users result in a bigger size of close friends' lists and vice versa. Medeiros and Bosse (2016) and Coursaris and Liu (2009) also expressed that empathetic abilities account for social support in social media, and Mayshak et al. (2017) showed that the level of user engagement with social networking

websites has a direct correlation with empathetic abilities. Finally, Del Rey et al. (2016) suggested that empathy negatively predicts traditional bullying and cyber-bullying perpetration.

In a health domain, recent studies show that empathy is one of the main advantages of using online health communities (OHCs) (Medeiros and Bosse, 2016; Nambisan, 2011; Malik and Coulson, 2010), which potentially fosters the healing process by decreasing distress and increasing optimism (Goubert et al., 2005; Olson, 1995). Table 1 shows an example from a cancer community, illustrating the function of empathetic messages.

The above studies, in social sciences and psychology, are based upon questionnaires, direct interviews, or at most hundreds of samples from manually collected data. These studies, however, suffer from several issues including scalability, biased data usage (Qiu et al., 2011), and high reliance on human memory that might not remember details accurately (Redelmeier and Kahneman, 1996; Litwin and McGuigan, 1999).

In the context of general social media, several computational studies started to analyze empathetic messages. However, these studies are contextually different from our study, which is focused on the health domain. For example, Rao et al. (2014) considered empathy as one of the eight classes of emotions in their classification task. In another work, Alam et al. (2017) annotated and modeled empathy in spoken conversations, based on multi-modal features extracted from conversations (such as acoustic features and video frames). As mentioned above, this is different from our work contextually and in terms of the applied methods. We use only textual comments.

Despite the importance of empathy in augmenting patients' positive feelings (Goubert et al., 2005; Olson, 1995), to our knowledge, there has not been any computational approaches proposed

| |
|---|
| *–Patient: Hi all sense being on chemo ( 5 down 1 to tch ) with the last two really I have had a problem with my BP being high. I am having a problem with my heart racing. At rest it may get down to 86. When my oncologist did the muga scan it went from 68 to 63. I have never had a problem with my heart at all. I'm Very nervous.*<br>*–Commentator: I had much the same problem while doing chemo, the last 2 or 3 rounds were the worst. Try not to worry to much! By the way I am the proud owner of 3 chihuahuas. Blessings to you...Alison*<br>*–Patient: Thanks so much I feel allot better now. I did talk to my Dr and he is giving me meds to lower the rate. I feel like I spend my time fighting side affects LOL. Thanks sisters. Take care all* |

Table 1: A sample of an empathetic message and its impact on patient's emotion

for identifying and analyzing empathetic messages in OHCs. Computational studies that analyzed OHCs data, have focused on analyzing emotional and informational support in patients' messages (Biyani et al., 2014; Wang et al., 2014; Qiu et al., 2011). Zhao et al. (2014, 2011) used the result of analyzing social support in OHCs for identifying influential users. However, these works address emotional support in general and do not focus on identifying empathetic messages.

In this paper, we propose a computational approach to analyzing large amounts of messages in OHCs and to automatically identifying messages that contain empathy. This first study on identifying empathetic messages in OHCs aims to make an appropriate foundation for further, deeper, and scalable studies and developing applications. Automatic empathetic message identifier can be used by OHCs' moderators for monitoring communities mental health, cyber-bullying and cyber-stalking detection, measuring the level of users engagement in communities (Mayshak et al., 2017), predicting users' position in online communities (Kardos et al., 2017), as well as the loneliness of users (Pamukçu and Meydan, 2010). Furthermore, such an application can be employed in measuring nursing skills (Yu and Kirk, 2008), measuring the quality of online counseling sessions, and assessing the quality of human-robot interactions (Fung et al., 2016; Leite et al., 2013).

Our contributions in this work are as follows:

1. We propose a machine learning model for identifying empathetic messages in OHCs. To our knowledge, this is the first work on automatically detecting empathy in OHCs.

2. We experimentally validate our empathy identification model on a manually annotated dataset generated from the Cancer Survivors' Network of the American Cancer Society.

3. We show that in general empathetic messages are correlated with a positive change in participants' sentiments.

## 2 Data Collection and Annotation

We randomly selected 225 comments from 21 discussion threads in the Lung Cancer discussion board in a Cancer Survivor' Network (CSN)[1]. Following Biyani et al. (2014), we selected messages (i.e., sentences in comments) with length greater than four words. We ended up with 1041 messages in total. We integrated our collected data with 1066 messages extracted from the breast cancer discussion board in CSN that was provided by Biyani et al. (2014).

The purpose of the annotation was to tag empathetic messages through which the message providers intended to show their empathy towards other people. Two annotators (graduate students) contributed to the task. They were asked to get familiarized with the concept of empathy by reading two studies (i.e., Collins (2014) and Decety and Jackson (2004)) during a week. After a group meeting between annotators and researchers to share and discuss their understanding of empathetic messages in the presence of two psychologists, the annotation task began in an iterative fashion similar to prior studies and guidelines (DMello, 2016; Fort et al., 2016; Shanahan et al., 2006) . In each round, 200 messages were assigned and annotators discussed disagreements with researchers; 100% inter-annotator agreement (IAA) was achieved after each round of discussions. We used Cohen's kappa for measuring IAA. After three initial rounds of annotations, the remaining data (1507 messages) were assigned to the annotators where they achieved 87% IAA. The last round of the assigned data was adjudicated by one of the authors. Table 3 provides the distribution of empathetic messages in the two datasets (breast cancer (B-dataset) and lung cancer (L-dataset)). As can be seen, B-dataset has significantly more empathetic messages than the L-dataset.

---

[1] https://csn.cancer.org

| Hyper-parameter Settings |
| --- |
| –LSTM: W2vec-S=150, LR= 0.001, L2reg=$1E-5$, Decay rate=0.7, Dropout=0.5, Layer=5, 1-Max pooling, Order= 3 |
| –ConvLST:W2vec-S=150, LR= 0.01, L2reg=$1E-5$, Decay rate=0.7, Dropout=0.5, Layer=2, 1-Max pooling, Order= 2 |
| –CNN: W2vec-S=150, LR= 0.1, L2reg=$1E-5$, Decay rate=0.5, Dropout=0.5, Layer=2, 1-Max pooling, FRS=(1,2,3), NF=64 |

Table 2: Hyperparameter settings for each model.

## 3 Model

| Dataset | Empathetic msgs. | Percentage(%) |
| --- | --- | --- |
| B-dataset | 494 out of 1066 | 46.3 |
| L-dataset | 295 out of 1041 | 28.3 |

Table 3: Statistics from the data collections.

In this section, we describe our proposed model for empathetic messages identification in OHCs.

**Problem Statement:** Given a message (i.e., a sentence in a comment) in an OHC, $S = \{W_1, W_2, \cdots, W_i, \cdots, W_n\}$ containing $n$ words, the task is to classify it as empathetic or not.

### 3.1 Word Representations

We use word embeddings with an embedding matrix $E_w \in R^{d_w \times V_m}$ where $d_w$ is the embedding dimension and $V_m$ is the word vocabulary size. We generate word embedding matrices by using the whole CSN collected data (i.e., users' comments from June 2000 to June 2012) of three different dimensions (i.e., 75, 150, 300). We use W2vector module in Gensim (Řehůřek and Sojka, 2010).

### 3.2 Model Description

The proposed model for classifying empathetic messages combines convolutional and LSTM (long short-term memory) networks (which we call ConvLSTM). Our ConvLSTM network takes word embeddings as input and creates a sequence of dense, real-valued vectors: $E = (e_1, e_2, \cdots, e_T)$. By applying multiple convolutional layers to $E$ and using pooling, we obtain a dynamic sequence of feature vectors: $F = (f_1, f_2, \cdots, f_n)$, which is fed into the LSTM. The output of the LSTM network is given to a softmax function to compute the predictive probabilities, $p(y = k|S)$, of each of the classes given a message $S$ (see Figure 1).

## 4 Experiments

In this section, we present our optimization process and the results of our model. We report precision, recall, and F-1 score, all macro-averaged across 10 folds in a cross-validation setting.
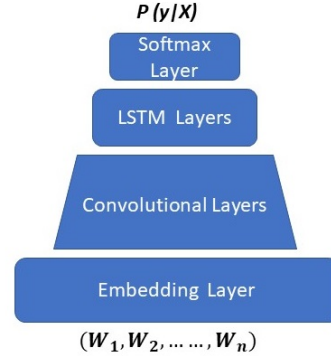


$P(y|X)$

Softmax Layer

LSTM Layers

Convolutional Layers

Embedding Layer

$(W_1, W_2, \ldots \ldots, W_n)$

Figure 1: ConvLSTM structure for empathetic message identification.

**Hyperparameters settings:** We optimized hyper-parameter values by performing a grid search on a development set, which consisted of 15% of instances in the training set in 10-fold cross validation experiments. We optimized hyperparameters of ConvLSTM and each of embedded models (i.e., CNN and LSTM) to compare their performances with ConvLSTM. We used a range of values for the following hyperparameters: word embedding vector size (i.e., $75, 150$, and $300$), learning rate (LR) $[0.1, 0.001]$, $l_2$ regularization (L2reg) $[0.0, 5E-5, 1E-5]$, decay rate $[0.0, 0.1, \cdots, 0.8]$, dropout $[0.0, 0.1, \cdots, 0.6]$, number of layers $[1, 2, \cdots, 10]$, pooling methods [1-Max, Mean, Last state], order size in LSTM {unigram, bigram, trigram}, filter region sizes (FRS) $[(1,2,3), (2,3,4), (3,4,5), (4,5,6)]$ and number of feature maps (NF) in CNN $[32, 64, \cdots, 256]$. Table 2 shows the best hyperparameters' settings by which each model achieved the best F-1 score on the development set.

**Baselines:** We compare our models with the following baselines:

1. **Bag-of-words and POS tags:** Word frequencies and their part-of-speech tags show the primary property of the text and has been used in studies on OHCs' message processing (Biyani et al., 2014, 2012b,a). We used both words and their POS tags' frequencies as features. We obtained the best performance using term-frequency encoding

248

| Method | P(%) | R(%) | F-1 (%) |
|---|---|---|---|
| ConvLST | 78.61 | **78.12** | **78.36** |
| LSTM | **79.47** | 75.00 | 77.17 |
| CNN | 76.20 | 77.00 | 76.60 |
| BoW+POS | 71.8 | 68.2 | 69.90 |
| Lexical-based | 54.5 | 46.9 | 50.4 |

Table 4: Empathetic message identification.

and document frequencies between 2 and 95% of the total documents. Multinomial Naïve Bayes achieved the best results among all evaluated classifiers (e.g., Support Vector Machines and Random Forest).

2. **Lexicon-based model:** Lexicon-based approaches have been used in many studies related to emotion detection (Strapparava and Mihalcea, 2007; Strapparava et al., 2004) and sentiment analysis tasks (Mohammad, 2012; Liu, 2012; Biyani et al., 2013). Following Biyani et al. (2014), we used the same lexicons. These lexicons include: weak and strong subjective words, cancer drugs, side-effects, and therapeutic procedures, for building our baseline's feature set.

**Empathetic Message Identification**: Table 4 compares the performance of our proposed model (ConvLSTM) with CNN, LSTM, and the baselines. As can be seen, our model achieves the best F-1 score, which is 8.46% higher than the F-1 score of the best baseline (i.e., 69.90%). Also, we can see that the combination of CNN and LSTM (ConvLSTM), which employs the sequences of important features extracted by CNN, achieves better performance than each of the individual CNN and LSTM models.

While LSTM achieved the best precision, ConvLSTM obtained the highest F-1 score and recall. Table 4 shows that Lexical-based baseline resulted in the lowest F-1 score. The lexicon-based baseline uses two types of features: subjectivity-related and informational-related features. After removing subjectivity features, the F-1 score drops to 15.7% and after removing informational features, the F-1 score drops to 47.3%. These results suggest that the subjectivity features are more effective than the informational-related ones, as expected, in identifying empathetic messages.

**Sentiment Dynamics with Empathetic Messages**: In this section, we conduct an experiment to investigate the potential of empathetic messages for changing the thread originator's feelings. We used the data extracted from CSN, which include
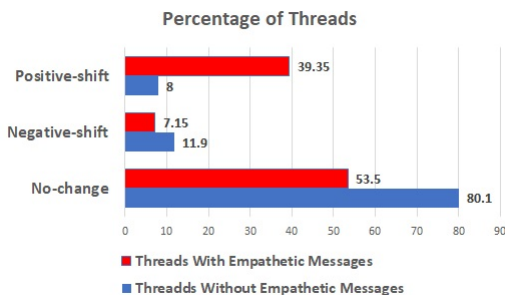


Figure 2: Thread-initiator's feeling transformation as a result of empathetic messages in a thread.

users' comments from June 2000 to June 2012. We extracted all threads where the originator of a thread replied (at least) once after an empathetic comment was posted from other users (responders). We followed the same experimental setting presented in Qiu et al. (2011). In total, 12915 discussion threads were extracted for analysis.

We ran our ConvLSTM model for empathetic message identification on all responders' messages, which were posted between two posts of the originator (e.g., the **Commentator**'s post in Table 1). We also discarded messages in which an initiator simply thanks a fellow member and used a threshold of four on the number of words (Biyani et al., 2014). We ran Stanford sentiment toolkit (Manning et al., 2014) on the originators' posts (e.g., the **Patient**'s posts in Table 1) to identify their sentiment. In this way, it is possible to determine whether the empathetic messages provided by responders who replied to the thread, are able to change the sentiment of the thread originator. To better understand any changes in feelings, we categorized changes in three groups, i.e., Positive-shift, Negative-shift, and No-change. Positive-shift represents any positive change in the sentiment of the thread initiator such as negative-to-positive, neutral-to-positive, negative-to-neutral. Negative-shift has a converse settings compared with the Positive-shift and No-change represents a state that originator's second post reflects the same sentiment as the initial one.

These results are shown in Figure 2 (the red bars). As can be seen from the figure, in 39.35% of the threads, empathetic messages bring a positive-shift in originators' feelings as opposed to only 7.15% negative-shift. We can also observe that in 53.5% of the threads, the originators' feelings do not change. Thus, we can conclude that empathetic messages play a major role in improving

participants' feelings in OHCs.

We also contrasted the positive-shift, negative-shift, and no-change in the threads with empathetic messages (the red bars in Figure 2) with those in the threads without empathetic messages (the blue bars in Figure 2) to better understand the impact of empathy on people's moods. More precisely, we ran the sentiment tool over the threads with no empathetic messages and found that only 8% positive shift, 11.9% negative shift and 80.1% no-change occurred. These results suggest that positive sentiment changes occur more prominently in threads containing empathetic messages compared to those with no empathetic messages.

## 5 Conclusion and Future Work

In this paper, we presented a machine learning model for identifying empathetic messages in online health communities. Our model is based on a combination of Convolutional Neural Networks and Long Short Term Memory networks, called ConvLSTM. We showed that ConvLSTM outperforms strong baselines. Moreover, we showed that empathetic messages do cause positive shifts in patients' sentiments in OHCs. In future, it would be interesting to investigate empathy identification in other sub-forums and the relation between the number of empathetic messages in a thread and the change in thread originators' emotional states.

## Acknowledgments

## References

Firoj Alam, Morena Danieli, and Giuseppe Riccardi. 2017. Annotating and modeling empathy in spoken conversations. *arXiv preprint arXiv:1705.04839*.

C Daniel Batson. 2009. These things called empathy: eight related but distinct phenomena.

Prakhar Biyani, Sumit Bhatia, Cornelia Caragea, and Prasenjit Mitra. 2012a. Thread specific features are helpful for identifying subjectivity orientation of online forum threads. In *COLING*, pages 295–310.

Prakhar Biyani, Cornelia Caragea, Prasenjit Mitra, and John Yen. 2014. Identifying emotional and informational support in online health communities. In *COLING*, pages 827–836.

Prakhar Biyani, Cornelia Caragea, Prasenjit Mitra, Chong Zhou, John Yen, Greta E Greer, and Kenneth Portier. 2013. Co-training over domain-independent and domain-dependent features for sentiment analysis of an online cancer support community. In *Proc. of the 2013 IEEE/ACM ASONAM*, pages 413–417.

Prakhar Biyani, Cornelia Caragea, Amit Singh, and Prasenjit Mitra. 2012b. I want what i need!: analyzing subjectivity of online forum threads. In *Proc. of the 21st ACM CIKM*, pages 2495–2498. ACM.

Franklin M Collins. 2014. The relationship between social media and empathy.

Constantinos K Coursaris and Ming Liu. 2009. An analysis of social support exchanges in online hiv/aids self-help groups. *Computers in Human Behavior*, 25(4):911–918.

M Davis, L Tiedens, and C Leach. 2004. Negotiating the border between self and other. *The social life of emotions*, pages 19–42.

Jean Decety and Philip L Jackson. 2004. The functional architecture of human empathy. *Behavioral and cognitive neuroscience reviews*, 3(2):71–100.

Rosario Del Rey, Lambros Lazuras, José A Casas, Vassilis Barkoukis, Rosario Ortega-Ruiz, and Haralambos Tsorbatzoudis. 2016. Does empathy predict (cyber) bullying perpetration, and how do age, gender and nationality affect this relationship? *Learning and Individual Differences*, 45:275–281.

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.

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

Pascale Fung, Dario Bertero, Yan Wan, Anik Dey, Ricky Ho Yin Chan, Farhad Bin Siddique, Yang Yang, Chien-Sheng Wu, and Ruixi Lin. 2016. Towards empathetic human-robot interactions. *arXiv preprint arXiv:1605.04072*.

Liesbet Goubert, Kenneth D Craig, Tine Vervoort, Stephen Morley, MJL Sullivan, Williams de CAC, A Cano, and Geert Crombez. 2005. Facing others in pain: the effects of empathy. *Pain*, 118(3):285–288.

Peter Kardos, Bernhard Leidner, Csaba Pléh, Péter Soltész, and Zsolt Unoka. 2017. Empathic people have more friends: Empathic abilities predict social network size and position in social network predicts empathic efforts. *Social Networks*, 50:1–5.

Jacques Launay, Eiluned Pearce, Rafael Wlodarski, Max van Duijn, James Carney, and Robin IM Dunbar. 2015. Higher-order mentalising and executive functioning. *Personality and individual differences*, 86:6–14.

Iolanda Leite, André Pereira, Samuel Mascarenhas, Carlos Martinho, Rui Prada, and Ana Paiva. 2013. The influence of empathy in human–robot relations. *International journal of human-computer studies*, 71(3):250–260.

Mark S Litwin and Kimberly A McGuigan. 1999. Accuracy of recall in health-related quality-of-life assessment among men treated for prostate cancer. *Journal of Clinical Oncology*, 17(9):2882–2882.

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

Sumaira H Malik and Neil S Coulson. 2010. Coping with infertility online: an examination of self-help mechanisms in an online infertility support group. *Patient education and counseling*, 81(2):315–318.

Christopher D Manning, Mihai Surdeanu, John Bauer, Jenny Rose Finkel, Steven Bethard, and David McClosky. 2014. The stanford corenlp natural language processing toolkit. In *ACL (System Demonstrations)*, pages 55–60.

Richelle Mayshak, Stefanie J Sharman, Lucy Zinkiewicz, and Alexa Hayley. 2017. The influence of empathy and self-presentation on engagement with social networking website posts. *Computers in Human Behavior*, 71:362–377.

Lenin Medeiros and Tibor Bosse. 2016. Empirical analysis of social support provided via social media. In *International Conference on Social Informatics*, pages 439–453. Springer.

Saif M Mohammad. 2012. # emotional tweets. In *Sixth International Workshop on Semantic Evaluation*, pages 246–255.

Priya Nambisan. 2011. Information seeking and social support in online health communities: impact on patients' perceived empathy. *Journal of the American Medical Informatics Association*, 18(3):298–304.

Joanne K Olson. 1995. Relationships between nurse-expressed empathy, patient-perceived empathy and patient distress. *Journal of Nursing Scholarship*, 27(4):317–322.

Burcu Pamukçu and Betül Meydan. 2010. The role of empathic tendency and perceived social support in predicting loneliness levels of college students. *Procedia-Social and Behavioral Sciences*, 5:905–909.

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 *PASSAT and SocialCom*, pages 274–281. IEEE.

Yanghui Rao, Qing Li, Liu Wenyin, Qingyuan Wu, and Xiaojun Quan. 2014. Affective topic model for social emotion detection. *Neural Networks*, 58:29–37.

Donald A Redelmeier and Daniel Kahneman. 1996. Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain*, 66(1):3–8.

Radim Řehůřek and Petr Sojka. 2010. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50.

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

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

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

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

Juping Yu and Maggie Kirk. 2008. Measurement of empathy in nursing research: systematic review. *Journal of Advanced Nursing*, 64(5):440–454.

Kang Zhao, Baojun Qiu, Cornelia Caragea, Dinghao Wu, Prasenjit Mitra, John Yen, Greta E. Greer, and Kenneth Portier. 2011. Identifying leaders in an online cancer survivor community. In *Proceedings of the 21st Annual Workshop on Information Technologies and Systems (WITS 2011)*.

Kang Zhao, John Yen, Greta Greer, Baojun Qiu, Prasenjit Mitra, and Kenneth Portier. 2014. Finding influential users of online health communities: a new metric based on sentiment influence. *Journal of the American Medical Informatics Association*, 21(e2):e212–e218.