*Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, pages 5444–5449
Marseille, 20-25 June 2022


# EnSyNet: A Dataset for Encouragement and Sympathy Detection

**Tiberiu Sosea** **Cornelia Caragea**
Computer Science
University of Illinois Chicago
tsosea2@uic.edu cornelia@uic.edu

**Abstract**

More and more people turn to Online Health Communities to seek social support during their illnesses. By interacting with peers with similar medical conditions, users feel emotionally and socially supported, which in turn leads to better adherence to therapy. Current studies in Online Health Communities focus only on the presence or absence of emotional support, while the available datasets are scarce or limited in terms of size. To enable development on emotional support detection, we introduce EnSyNet, a dataset of $6,500$ sentences annotated with two types of support: encouragement and sympathy. We train BERT-based classifiers on this dataset, and apply our best BERT model in two large scale experiments. The results of these experiments show that receiving encouragements or sympathy improves users' emotional state, while the lack of emotional support negatively impacts patients' emotional state. We make our code and data available at https://github.com/tsosea2/EnSyNet.

**Keywords:** Sympathy, Encouragement, Online Health Communities

## 1. Introduction

Online Health Communities (OHCs) have become a major source of support for people with health problems and for their caregivers. People usually seek social support during their treatments (LaCoursiere, 2001), and such support is a key component to their overall well-being, having the potential to directly affect the success of their treatment (Coulson et al., 2007; Kim et al., 2004). Social support in OHCs can be divided in two core components: *emotional* and *informational* support (Olson, 1995; Gooden and Winefield, 2007; Eysenbach et al., 2004; Brennan and Ripich, 1994). These two core functions of OHCs help to promote new treatments, lifestyle, and can also reveal adverse drug effects (Yang et al., 2012).

In this paper, we focus on emotional support detection, more specifically on the expression of encouragements and sympathy. Encouragement is *the action of giving someone support, confidence, or hope*, while sympathy is *the feeling of pity or sorrow for someone else's misfortune*. These two types of support are extremely beneficial for patients in improving their emotional state. To our knowledge, we are the first to address this fine-grained approach, as previous studies on emotional support detection focused only on a high level view of emotional support, which analyzes the presence or absence of such support (Khanpour et al., 2018; Biyani et al., 2014; Wang et al., 2014). To this end, we introduce EnSyNet, a new dataset composed of $6,500$ sentences annotated with encouragement and sympathy. EnSyNet is a challenging benchmark, being more fine-grained and 6 times larger than other datasets used for detecting emotional support (Biyani et al., 2014; Wang et al., 2014). To indicate a few challenges of the task, we show some examples from our dataset in Table 1. For instance, we observe that a sentence can contain both sympathy and encouragement. In *I know how terrible this damn disease is but keep the battle and I know I and many others are praying hard for you*, the encouragement is expressed through the expression *keep the battle*, while *praying hard* expresses sympathy towards the person. Surprisingly, distinguishing between sympathy and encouragement is challenging, and $40\%$ of the sentences that convey sympathy also express encouragements.

| | |
|---|---|
| SYM ENC | I know how terrible this damn disease is but keep the battle and I know I and many others are praying hard for you. |
| SYM | Sending lots of prayers and positive thoughts to you, my friend. |
| ENC | I wish your mom the best in making a decision. |
| NONE | Plus the chemo makes you lose your appetite anyway. |

Table 1: Examples from the dataset.

We model the expression of sympathy and encouragement in EnSyNet using weak and strong baselines. Notably, our best pre-trained language model achieves as much as $0.63$ F1 on sympathy and $0.75$ F1 on encouragement. We improve upon these results by using additional unsupervised pretraining on health-related and emotion-related datasets. Finally, we perform a comprehensive analysis of our dataset, then use models trained on EnSyNet in two large-scale experiments to understand the impact of sympathy and encouragement to the emotional states of patients.

Our contributions are as follows: **(1)** We create En-

SyNet, a new dataset of $6,500$ sentences for encouragement and sympathy detection. **(2)** We show EnSyNet is a challenging benchmark by fine-tuning state-of-the-art pre-trained language models, and exploring additional domain-specific pre-training for inductive biases. **(3)** We apply our best pre-trained language model in a large-scale experiment across the OHC. Our experiments validate the importance of social support in OHCs, revealing that offering sympathy or encouragement to users improve their state-of-mind, increasing the emotion frequency of joy, while decreasing that of sadness.

## 2. Related Work

Despite that a wide body of research focuses on analyzing social support in Online Health Communities using surveys and questionnaires (Coursaris and Liu, 2009; Han et al., 2011; Zaphiris and Pfeil, 2007; Rodgers and Chen, 2005; Buis, 2008; Coulson et al., 2007; Kim et al., 2004), computational studies of social support started to emerge only recently. Yang et al. (2019), Wang et al. (2015) and Nolker and Zhou (2005), modeled social roles in OHCs, by identifying clearly defined roles such as *emotional support seeker* or *informational support provider*. Wang et al. (2012) studied how social support in OHCs affects the length of user participation, while Chancellor et al. (2018) studied behavioral changes in the presence of social support. Wang et al. (2014), Biyani et al. (2014), and Khanpour et al. (2018) studied emotional support at a high level, by only identifying the presence or degree of emotional support. Instead, in order to get a better understanding into how emotional support is provided across OHCs, we adopt a more fine-grained approach by studying 2 specific types of emotional support: encouragements and sympathy.

## 3. Dataset

We build our dataset from Cancer Survivors Network[1] (CSN), which is one of the largest online health communities for people suffering from cancer. Here, users can start discussion threads or reply to existing discussions, and these threads are grouped by cancer type. We sample our data from three frequent types of cancer: breast, lung and prostate cancer. We choose to model the task at sentence level, because consecutive sentences in a discussion might cover different topics (Biyani et al., 2014). The **task objective** is to predict if a sentence contains sympathy or encouragement.

**Construction** We randomly sampled $25,000$ sentences from CSN and used the Amazon Mechanical Turk crowd-sourcing platform to acquire labels for the data. For every sentence, the final label can be sympathy, encouragement, both sympathy and encouragement, or neither. We designed separate annotation forms for sympathy and encouragement. These forms ask an annotator to label a sentence with *true* or *false*, i.e., if a sentence contains sympathy/encouragement, the label is *true*, otherwise it is *false*. We use 5 annotators for the labelling, and remove spurious annotations by ruling out annotators inconsistent with the majority vote in more than $20\%$ of the cases. This setting provides a good inter-annotator agreement, with a Krippendorff $\alpha = 0.71$.

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

| | SYM | ENC |
|---|---|---|
| SYM | $2,595$ | $1,004$ |
| ENC | $1,004$ | $4,881$ |

Table 2: Co-occurence of sympathy and encouragement.

**Analysis** We present the frequency and co-occurence of sympathy and encouragement in Table 2. Interestingly, there are about twice more messages containing encouragement compared to sympathy. We also observe that $\sim 40\%$ of sentences expressing sympathy also contain encouragement. Finally, out of the $25,000$ sentences sampled, $25\%$ contain sympathy or encouragement.

**Benchmark Dataset** We provide a benchmark dataset created as follows. For each of the encouragement and sympathy categories, we consider as *positive* examples all sentences containing encouragement or sympathy, respectively. Next, we sample an equal number of *negative* examples in the following manner: **(1)** For encouragement, we use all the sentences conveying sympathy and not encouragement, while we sample the rest from the set of sentences with neither sympathy or encouragement. **(2)** For sympathy, $\frac{1}{3}$ are sampled from sentences conveying neither sympathy or encouragement, while $\frac{2}{3}$ are sampled from sentences containing encouragement but not sympathy. Finally, we randomly construct a 80/10/10 train, validation and test split, which we will provide alongside our data.

## 4. Experiments and Results

### 4.1. Methods

We evaluate the sympathy and encouragement detection performance on EnSyNet using several methods.

**Weak Baselines** We experiment with **1)** Naive Bayes applied after stemming and stop-word removal and **2)** Logistic Regression using tf*idf vectors as input.

**Strong Baselines** As strong baselines, we explore with: **1)** Bi-directional Long Short Term Memory network (Bi-LSTM). (Hochreiter and Schmidhuber, 1997), **2)** Convolutional Neural Network (CNN) (Kim, 2014) using pre-trained GloVe (Pennington et al., 2014) word embeddings, and **3)** Convolutional Bi-LSTM (Conv-Bi-LSTM) (Khanpour and Caragea, 2018) using pre-trained GloVe (Pennington et al., 2014) word embeddings.

| METHOD | SYMPATHY | | | ENCOURAGEMENT | | |
|---|---|---|---|---|---|---|
| | PRECISION | RECALL | F-1 | PRECISION | RECALL | F-1 |
| NAIVE BAYES | 0.53 | 0.55 | 0.54 | 0.70 | 0.74 | 0.67 |
| LOG REGRESSION | 0.57 | 0.58 | 0.58 | 0.66 | 0.68 | 0.67 |
| BI-LSTM | 0.58 | 0.58 | 0.58 | 0.67 | 0.72 | 0.69 |
| CNN | 0.58 | 0.60 | 0.59 | 0.68 | 0.70 | 0.69 |
| CONV-BI-LSTM | 0.56 | 0.57 | 0.57 | 0.67 | 0.67 | 0.67 |
| BERT | 0.61 | 0.63 | 0.62 | 0.74 | 0.72 | 0.73 |
| DISTILBERT | 0.62 | 0.64 | 0.63 | 0.76 | 0.74 | 0.75 |
| CLINICAL BERT | 0.60 | 0.62 | 0.61 | 0.74 | 0.70 | 0.72 |
| DISTILBERT+EMONET | 0.57 | 0.59 | 0.58 | 0.64 | 0.62 | 0.63 |
| DISTILBERT+CSN | 0.62 | 0.67 | 0.64 | 0.79 | 0.74 | 0.76 |

Table 3: Precision, recall and F-1 results for sympathy and encouragement detection. We show results obtained using weak baselines (first block), strong baselines (second block), pre-trained language models (third block), and language models that underwent additional pre-training using masked language modeling (fourth block).

**Pre-trained Language Models** Finally, we experiment with pre-trained language models that use transfer learning, a method which implies pre-training on a large unsupervised task, followed by fine-tuning on a less computationally expensive supervised task. We experiment with **1)** BERT (Devlin et al., 2018) base uncased model, **2)** DistilBERT (Sanh et al., 2019), a distilled version of BERT with a considerably reduced size, and **3)** Clinical BERT (Alsentzer et al., 2019), a specialized version of BERT model pre-trained on a health-related task (e.g., clinical narratives, medical text.

**Additional Pre-training** It has been shown that performing additional unsupervised pretraining before the fine-tuning phase can prove successful when dealing with specialized contexts (Sosea and Caragea, 2020; Desai et al., 2020; Gururangan et al., 2020) (such as medical forums). To this end, to improve the performance of our best language models, we train our model further on various pre-training domains using dynamic masked language modelling (Devlin et al., 2018). We consider two pre-training domains. First, we pre-train our model on EmoNet (Abdul-Mageed and Ungar, 2017), a dataset annotated with fine-grained emotions using distant supervision, obtained from the authors. Given that the EmoNet sentences are biased towards an emotion context, which is somewhat close to our emotional support area, we investigate if pre-training on this task improves the model performance. Second, we pre-train on the entire CSN forum, hoping to implicitly induce health-specific biases.

## 4.2. Experimental Setup

We carry out all our experiments using an Nvidia V100 GPU. We use the HuggingFace Transformers (Wolf et al., 2020) library for the implementation of our pre-trained language models. We run all our experiments on the benchmark dataset 10 times using different weight initializations, and report the average of the runs.


(a) Encouragement (b) Sympathy

Figure 1: Word Clouds from sentences containing encouragement or sympathy

## 4.3. Results

We show the results of the above models in Table 3. First, the pre-trained language models consistently outperform the other baselines by as much as 4% on sympathy and 6% on encouragement. The DistilBERT model obtains the best performance among the language models, even though it is almost half the size and a direct distilled version of BERT. Second, the pre-trained models tend to have a better precision than recall on encouragement, and opposite on sympathy. We also observe that Clinical BERT does not help the performance on the task, even though it is specifically pre-trained on health-related data. Finally, additional pre-training on CSN improves the downstream overall performance on both sympathy and encouragement, whereas pre-training on EmoNet (Abdul-Mageed and Ungar, 2017) largely hurts the performance. To investigate why EmoNet pre-training decreases the downstream performance, we perform a linguistic analysis. First, EmoNet is constructed using distant supervision

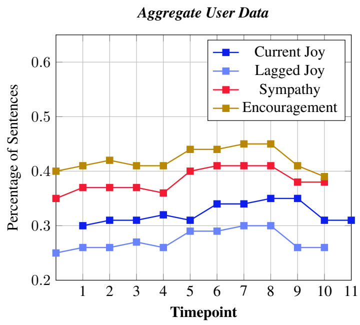


Figure 2: Received sympathy and encouragement, as well as joy level of a user in time.

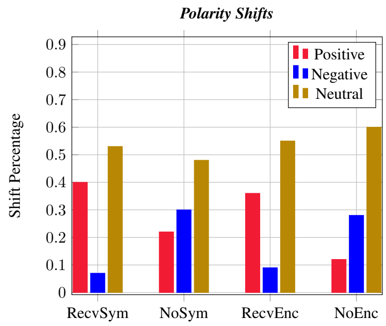


Figure 3: Shifts in feeling polarity correlated with receiving or the lack of sympathy and encouragement.

from a Twitter corpus by leveraging cues in the data in the form of emotion hashtags. This domain differs substantially in language and terminology from our health-related domain. Second, we perform a word-level analysis and extract the most prominent words in our data, and show the results using word clouds in Figure 1a and 1b. Interestingly, we observe a notable difference between the word sets used in our sympathy and encouragement sentences, which are not related to the emotional context that EmoNet (Abdul-Mageed and Ungar, 2017) addresses.

## 5. Large-Scale Experiments

To illustrate the importance of encouragement and sympathy in Online Health Communities we apply our best performing model in two large-scale experiments. In these experiments, we investigate correlations between sympathy/encouragement and emotions such as joy and sadness in order to study if emotional support helps improve a patient's feelings and state-of-mind. We obtained models for joy and sadness sentence classification trained on CSN from the authors of a previous work on fine-grained emotion detection (Khanpour and Caragea, 2018).

**Sympathy and Encouragement effect in Time** We randomly sample 20 users from the entire CSN forum with at least 500 posts. For each user, we sort the posts chronologically, then group them in 11 timepoints with an equal number of posts in each timepoint. For each timepoint $t$ and an user $u$, we look at all the comments addressed to $u$ and use our models to calculate the percentage of sentences containing encouragements or sympathy. Next, for timestep $t+1$, we extract all of $u$'s posts and, using the trained emotion model, we determine the percentage of sentences expressing joy. We do this for all the 11 timepoints and for all 20 users, and average the results. We now address the following question: *Does encouragement and sympathy offered to an user at timepoint t correlate with the joy of that same user at a later time $t+1$?*. We show our findings in Figure 2. For visualization purposes, we also show a *Lagged Joy* graph. This graph is the same as the Joy graph, but translated 1 timepoint backwards on the $x$ axis, and $0.05$ downwards on the $y$ axis. We perform these operations for better visualization and to get an alignment in timesteps with the encouragement and sympathy. In our experiments, $t$ is one month. Interestingly, the results show a strong correlation between joy and the amount of emotional support received: emotional support boosts a user's joy, while the absence of such support tends to worsen the feelings of the user.

**Potential of Emotional Support to Change Feelings** In this experiment, we investigate how likely messages conveying emotional support are to change a user's feelings. To this end, we sample 5000 threads from the CSN with the following property: let $x$ be the post that initiates the thread. There has to be a post $y$ somewhere in the thread, such that the author of $x$ is the same as the author of $y$. First, we compute the polarity of post $x$. The polarity is positive if there are more sentences conveying joy than sadness, neutral if they are equal, and negative if there are more sentences expressing sadness. Then, we compute the polarity of post $y$. Next, we check if the messages inbetween these two posts contain sympathy or encouragement. Finally, we investigate if positive shifts in polarity happen more frequently if the user receives encouragement or sympathy, as well as if negative polarity shifts appear more often in the absence of encouragement or sympathy. We show our results in Figure 3. Interestingly, there are significantly more positive shifts and less negative shifts if users are offered emotional support. For instance, the probability of a user to have a negative polarity shift is four times less likely if being offered sympathy than not.

## 6. Conclusion

We introduced EnSyNet, a new health related dataset of $6,500$ sentences aimed at detecting two types of emotional support: sympathy and encouragement. Our experiments show that EnSyNet is a challenging benchmark, even for large-scale pre-trained language models, obtaining an F-1 of $76\%$ on encouragement detection and $64\%$ on sympathy detection. We make our code and dataset publicly available.

## Acknowledgments

This work is partially supported by the NSF Grants IIS-2107487 and IIS-1912887. Any opinions, findings, and conclusions expressed here are those of the authors and do not necessarily reflect the views of NSF. The computation for this project was performed on Amazon Web Services through a research grant. We thank our reviewers for their feedback and comments.

## Bibliographical References

Abdul-Mageed, M. and Ungar, L. (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.

Alsentzer, E., Murphy, J., Boag, W., Weng, W.-H., Jindi, D., Naumann, T., and McDermott, M. (2019). Publicly available clinical BERT embeddings. In *Proceedings of the 2nd Clinical Natural Language Processing Workshop*, pages 72–78, Minneapolis, Minnesota, USA, June. Association for Computational Linguistics.

Biyani, P., Caragea, C., Mitra, P., and Yen, J. (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.

Brennan, P. F. and Ripich, S. (1994). Use of a home-care computer network by persons with aids. *International journal of technology assessment in health care*, 10(2):258–272.

Buis, L. R. (2008). Emotional and informational support messages in an online hospice support community. *CIN: Computers, Informatics, Nursing*, 26(6):358–367.

Chancellor, S., Hu, A., and De Choudhury, M. (2018). Norms matter: contrasting social support around behavior change in online weight loss communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–14.

Coulson, N. S., Buchanan, H., and Aubeeluck, A. (2007). Social support in cyberspace: a content analysis of communication within a huntington's disease online support group. *Patient education and counseling*, 68(2):173–178.

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

Desai, S., Caragea, C., and Li, J. J. (2020). Detecting perceived emotions in hurricane disasters. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5290–5305, Online, July. Association for Computational Linguistics.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., and Stern, A. (2004). Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *Bmj*, 328(7449):1166.

Gooden, R. J. and Winefield, H. R. (2007). Breast and prostate cancer online discussion boards: a thematic analysis of gender differences and similarities. *Journal of Health Psychology*, 12(1):103–114.

Gururangan, S., Marasović, A., Swayamdipta, S., Lo, K., Beltagy, I., Downey, D., and Smith, N. A. (2020). Don't stop pretraining: Adapt language models to domains and tasks. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8342–8360, Online, July. Association for Computational Linguistics.

Han, J. Y., Shah, D. V., Kim, E., Namkoong, K., Lee, S.-Y., Moon, T. J., Cleland, R., Bu, Q. L., McTavish, F. M., and Gustafson, D. H. (2011). Empathic exchanges in online cancer support groups: distinguishing message expression and reception effects. *Health communication*, 26(2):185–197.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.

Khanpour, H. and Caragea, C. (2018). Fine-grained emotion detection in health-related online posts. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1160–1166.

Khanpour, H., Caragea, C., and Biyani, P. (2018). Identifying emotional support in online health communities. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), Apr.

Kim, S. S., Kaplowitz, S., and Johnston, M. V. (2004). The effects of physician empathy on patient satisfaction and compliance. *Evaluation & the health professions*, 27(3):237–251.

Kim, Y. (2014). Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*.

LaCoursiere, S. P. (2001). A theory of online social support. *Advances in Nursing Science*, 24(1):60–77.

Nolker, R. D. and Zhou, L. (2005). Social computing and weighting to identify member roles in online communities. In *The 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI'05)*, pages 87–93. IEEE.

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

Pennington, J., Socher, R., and Manning, C. (2014). GloVe: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, Doha, Qatar, October. Association for Computational Linguistics.

Rodgers, S. and Chen, Q. (2005). Internet community group participation: Psychosocial benefits for women with breast cancer. *Journal of Computer-Mediated Communication*, 10(4):JCMC1047.

Sanh, V., Debut, L., Chaumond, J., and Wolf, T. (2019). Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*.

Sosea, T. and Caragea, C. (2020). CancerEmo: A dataset for fine-grained emotion detection. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8892–8904, Online, November. Association for Computational Linguistics.

Wang, Y.-C., Kraut, R., and Levine, J. M. (2012). To stay or leave? the relationship of emotional and informational support to commitment in online health support groups. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*, pages 833–842.

Wang, X., Zhao, K., and Street, N. (2014). Social support and user engagement in online health communities. In *International Conference on Smart Health*, pages 97–110. Springer.

Wang, X., Zuo, Z., and Zhao, K. (2015). The evolution of user roles in online health communities-a social support perspective. In *PACIS*, page 121.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., Drame, M., Lhoest, Q., and Rush, A. (2020). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online, October. Association for Computational Linguistics.

Yang, C. C., Yang, H., Jiang, L., and Zhang, M. (2012). Social media mining for drug safety signal detection. In *Proceedings of the 2012 international workshop on Smart health and wellbeing*, pages 33–40.

Yang, D., Kraut, R. E., Smith, T., Mayfield, E., and Jurafsky, D. (2019). Seekers, providers, welcomers, and storytellers: Modeling social roles in online health communities. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–14.

Zaphiris, P. and Pfeil, U. (2007). Introduction to social network analysis. In *Proceedings of HCI 2007 The 21st British HCI Group Annual Conference University of Lancaster, UK 21*, pages 1–2.