A report of the CL-Aff OffMyChest Shared Task: Modeling Supportiveness and Disclosure

Kokil Jaidka[1], Iknoor Singh[2], Jiahui Lu[3], Niyati Chhaya[4], and Lyle Ungar[5]

[1] National University of Singapore, Singapore
[2] Panjab University, India
[3] Nanyang Technological University, Singapore
[4] Adobe Research, India
[5] University of Pennsylvania, USA
jaidka@nus.edu.sg

Abstract. This overview describes the official results of the CL-Aff Shared Task 2020 – #OffMyChest. The dataset comprised a semi-supervised classification task, and an open-ended knowledge modeling task on a dataset of Reddit comments with annotations crowdsourced from Amazon Mechanical Turk. The Shared Task was organized as a part of the 3^{rd} Workshop on Affective Content Analysis @ AAAAI-20, held in New York, USA, on February 7, 2020. This paper compares the participating systems in terms of their accuracy and F-1 scores at predicting different facets of self-disclosure. Feedback from the system runs was used to weed out labeling errors in the test set. The annotated test and training datasets, instructions, and the scripts used for evaluation are available at the GitHub repository.

1 Introduction

There is a growing interest in understanding how humans initiate and hold conversations online. A plethora of social media platforms has emerged and been adopted by internet communities worldwide. Different cultures and communities have emerged around different social media platforms [3], where some social networking sites are intended more for discussions among professional contacts, e.g., LinkedIn; others are often appropriate for pursing topical interests, e.g., Twitter; for having reasoned debates, e.g., Reddit; still others were developed to provide technical support, e.g., StackOverflow. A defining feature of these platforms is how their social norms differ. On different platforms, people choose to respond differently to each other and share different kinds of information about themselves [6]. An interesting research problem that arises is to quantify the levels of disclosure and to apply them for cross-sectional or longitudinal analysis of social norms and platforms. In this Shared Task, we take the first step towards approaching these problems, by examining the affective aspect of online conversations among strangers. Our aim is to build a new resource to model how social media users reciprocate in conversations, with emotional and informational behavior that either offers self-revelation or moral support. In this

 In: N. Chhaya, K. Jaidka, J. Healey, L. H. Ungar, A. Sinha (eds.): Proceedings of the 3rd Workshop of Affective Content Analysis, New York, USA, 07-FEB-2020, published at http://ceur-ws.org

paper, we introduce the OffMyChest conversation dataset and present the results of the concluded 2nd Computational Linguistics Affect Understanding (CL-Aff) Shared Task on modeling interactive affective responses. It was held in February 2020 as a part of the AAAI Annual Meeting in New York.

2 Background

Previous work exploring disclosure and support has usually examined its evidence in health forums [14,12]. In studies on general social media posts [11], women were found to self-disclose more than men, and people with a stronger desire for impression management are less likely to disclose about themselves online. Cross-platform differences in language can enable greater or lesser predictive accuracy at identifying users' demographic information [6]. Anonymity is one of the many technological affordances which is expected to make it easier for individuals to express negative feelings online [8]. Previous findings offer a way to understand how platform behavior can differ, but they do not differentiate between the information and emotional aspects of disclosure and support. Our Shared Task is motivated to address this research gap and to offer a way to distinguish emotional expressions from emotional support and informational disclosure from informational support. The ability to distinguish between these aspects would allow targeted interventions where mental health issues may be evident or where users' personal information may be at risk when they share too many personal details about themselves.

The work closest to our interest has provided annotation schemes to codify the type of disclosure [2] and support [12] in online help forums. Their work reports that support forums offer a higher degree of self-disclosure than discussion forums [2]. Furthermore, they reported that self-disclosure was often reciprocal, and reciprocity was more likely among female than male respondents. Other findings suggest that it is emotional support [12], rather than information support, that predicts users' longevity in a health support group. On the other hand, informational support satisfied members' short-term information needs.

We were inspired to explore how easily these notions of disclosure and support can generalize into understanding casual conversations between users. To denoise the data, we decided to focus on discussions of relationships and opted to focus on Reddit sub-communities, which are likely to offer better training data thanks to the enforced community rules and strict moderation.

First, we provide the definitional scope of disclosure and support for the CL-Aff Shared Task:

Emotional Disclosure: Comments that mention the author's feelings. Examples:

– "My only concern was for my son."
– "Fuck me that is beautiful."
– "Thanks for sharing the story."
– "My heart melted reading this xx";
– "I'm literally too jealous";

– "My heart is breaking for you."

Informational disclosure: Comments that contain at least some personal information about the author. Examples:

– "I'm now 65 years old";
– "I've worked with kids with ODD and autism."
– "I live in West Philly."
– "Sounds like our bipolar kid.";
– "She posted a screenshot of his porn history (gross)";
– "My mum told me that she was sexually abused as a kid."

Emotional Support: The comment is offering sympathy, caring, or encouragement. Examples:

– "Good luck, this shit is tough";
– "Good luck! but I'm afraid I have no advice";
– "You sound like a great person";
– "I'm so sorry.";
– "That's a great story."

Informational support: This comment is offering specific information, practical advice, or suggesting a course of action. Examples:

– "I wouldnt..";
– "You shouldn't..";
– "You can't..".;
– "Why didn't you try this?";
– "Please talk to a professional."

3 Corpus

On Reddit, discussions of relationships typically happen on the r/relationships community. However, a preliminary examination suggested that the discussions are not the kind of 'casual' conversations we were aiming for, and are instead more similar to a support forum. Responses to posts in this community would be skewed towards greater support and disclosure. We wanted a neutral, easy-to-generalize situation, where the pressure to reciprocate is substantively reduced. After further exploration, we decided to mix data from two subreddits. The first one we selected was r/CasualConversations, a 'friendlier' sub-community where people are encouraged to share what's on their mind about any topic. In essence, this is similar to the posting behavior encouraged on a typical social media platform. The second one we selected was r/OffmyChest, intended as 'a mutually supportive community where deeply emotional things you can't tell people you know can be told.' We anticipated that a mixture of labeled data from both these platforms would give us a degree of heterogeneity in the confessional and emotional behavior while preserving the high topicality and post quality that is typical of Reddit posts. We provide further details of the dataset in the following subsections.

3.1 Dataset description

The CL-Aff corpus comprises the following:

- **Unlabeled training set of posts (N=17,392):** The top posts in 2018 in */r/CasualConversations* and */r/OffMyChest* mentioning any of the terms *boyfriend, girlfriend, husband, wife, gf, bf*. Posts that are parents of comments in the training and test sets are separately identified.
- **Unlabeled training set of comments (N = 420,000):** Over 420k sentences extracted from 130k comments posted to the unlabeled set of posts mentioned above.
- **Labeled training set (N = 12,860):** 12,860 labeled sentences, extracted from the top comments posted to the top posts of the Reddit communities mentioned above.
- **Test set: (N = 5,000)** Labeled sentences, extracted from the top comments made to the posts mentioned above.

A detailed breakdown of the labeled training and test sets is provided in Table 1.

Table 1: CL-Aff #OffmyChest dataset statistics. Total number of instances and positive instances for each of the labels provided.

	r/OffMyChest	**r/CasualConversation**
	Training set	
Emotional disclosure	2449	1499
Information disclosure	2749	2142
Emotional support	901	349
Information support	772	234
Total observations	**7613**	**5247**
	Test set	
Emotional disclosure	2301	1237
Information disclosure	1237	1158
Emotional support	1094	406
Information support	854	316
Total observations	**3257**	**1743**

3.2 Data collection

Data was collected by first subsetting on the posts discussing relationships that were posted to either *r/OffmyChest* or *r/CasualConversation*. Posts about relationships were identified based on the presence of the seed words relating to romantic partners. Posts were then deduplicated, and all their underlying comments were collected. A sentence splitter was applied to obtain sentences, and a random sample of sentences which were at least 10 characters in length was then used for the pilot and confirmatory annotation tasks.

4 Annotation

Annotators were required to annotate each moment according to the inset questionnaire. The Disclosure and Support characteristics of each sentence were finally transformed into a binary (yes/no) coding and the labels were assigned based on a simple majority agreement between five independent annotators. Only labels with 60% - 100% agreement were retained. The pairwise percentage agreement on the final dataset was 71.2% each for emotional and informational disclosure, and 84.5% and 83.9% for emotional and informational support.

Instructions In this job, you will be presented with a comment made on Reddit, a popular discussion forum worldwide. The topic of the discussion is a casual conversation or a confession. Review the text of the comment and help us by answering a few yes/no questions about it. Each HIT takes about 30 seconds:

<***Comment appears here***>
Is this comment SHARING PERSONAL FEELINGS? NO/A LITTLE/A LOT

- NO: This comment does not mention the author's feelings about anything. ("It's a book by Hemingway"; "Are you ok?"; "She was really mad at me.")
- A LITTLE: This comment mentions the author's mild positive or negative feelings. ("My only concern was for my son."; "Fuck me that is beautiful."; "Thanks for sharing the story.")
- A LOT: This comment contains deep positive or negative feelings or tears. ("My heart melted reading this xx"; "I'm not crying, you're crying!"; "I'm literally too jealous"; "My heart is breaking for you.")

<***Comment appears here***>
Is this comment SHARING PERSONAL INFORMATION? NO/A LITTLE/A LOT

- NO: This comment does not mention the author's feelings about anything. ("It's a book by Hemingway"; "Are you ok?"; "She was really mad at me.")
- A LITTLE: This comment mentions the author's mild positive or negative feelings. ("My only concern was for my son."; "Fuck me that is beautiful."; "Thanks for sharing the story.")
- A LOT: This comment contains deep positive or negative feelings or tears. ("My heart melted reading this xx"; "I'm not crying, you're crying!"; "I'm literally too jealous"; "My heart is breaking for you.")

<***Comment appears here***>
Is this comment SUPPORTIVE? YES/NO

- YES: This comment is offering support to someone, either through sympathy, encouragement, or advice. ("Good luck, this shit is tough"; "Good luck! but I'm afraid I have no advice"; "Hey you tried your best"; "Have you tried family therapy?")
- NO: This comment does not offer any support.. ("Thank you for your time."; "This is so sweet."; "Badass grandpa."; I'm now 65 years old"; "I've worked with kids with ODD and autism"; "I live in West Philly.")

<***Comment appears here***>
Is this comment SUPPORTIVE? YES/NO

– GENERAL SUPPORT: The comment is offering general support through quotes and catchphrases. ("What's the worst that could happen?"; "You only die once."; "All's well that ends well.") (YES/NO)
– INFORMATIONAL SUPPORT: The sentence is offering information, advice, or suggesting a course of action. ("I wouldnt.."; "You shouldn't.."; "You can't..". "Why didn't you try this?"; "Please talk to a professional.")
– EMOTIONAL SUPPORT: The sentence is offering sympathy, caring, or encouragement. ("Good luck, this shit is tough"; "Good luck! but I'm afraid I have no advice"; "You sound like a great person"; "I'm so sorry."; "That's a great story.")

5 Overview of Approaches

Twelve teams signed up, and six teams finally submitted their results by the Shared Task deadline. The following paragraphs discuss the approaches followed by the participating systems, sorted in alphabetical order:

– GATech USA[4]: The team from GATech followed a semi-supervised approach comprising transformer-based models. Their regularization was predicated on the assumption that the class distribution in the test set would be similar to that of the training set.
– Gyrfalcon[10]: The team from Gyrfalcon Technology, California, proposed an algorithm to map English words into squared glyphs images, which they call Super Characters. These were implemented on a CNN Domain-Specific Accelerator in order to capture properties of disclosure and support.
– International Institute of Information Technology India [9]: The IIIT-H team employed a predictive ensemble model that combined predictions from multiple models based on fine-tuned contextualized word embeddings, RoBERTa and ALBERT.
– Pennsylvania State University USA (PennState)[1]: The PennState team also followed an ensemble approach, but with BERT, LSTM, and CNN neural networks. In their first model, they performed classification using BERT, fine-tuned their word representations, and obtained the hidden attention and sentence representation features in the CNN model, where they replaced the typical embedding layer with the pre-trained BERT model.
– Sungkyunkwan team (SKKU)[5]: The SKKU team used a semi-supervised approach, with the original posts as contextual information, and applied BERT, GLoVe, and Emotional GLoVe embedding models, to represent the text for label prediction.
– University of Ottawa (UOttawa) Canada[13]: The University of Ottawa team applied a deep multi-task learning approach that employed the logical relationship among the different labels to create 'fragment layers,' that were used to build a multi-task deep neural network.

6 Results

6.1 Task 1: Predicting Disclosure and Support

This section compares the participating systems in terms of their performance. The results with the best-performing system runs from each of the participating teams are provided in Figure 1. The performance of individual system runs is provided in Table 2 and Table 3. For the detailed implementation of the individual runs, please refer to the system papers which are included in this proceedings volume.

Figure 1a shows that predicting disclosure was evidently a harder problem than predicting support. The best performance at predicting both emotional and informational disclosure was obtained from the team from UOttawa [13](Accuracy = .69). The second and third spots for predicting emotional disclosure went to IIIT [9] and GATech [4], with an accuracy of .62 and .61, respectively. Predictive performances for informational disclosure were rather close to one another, with Gyrfalcon [10] and GATech [4] coming in a close second- and third-places with accuracies of .64 and .63 respectively.

Figure 1b shows that IIIT [9], UOttawa [13], and GATech [4] were neck-and-neck at predicting emotional and informational support, with IIIT getting a slight edge thanks to its performance on emotional support.

The most successful runs can be identified by referring to Table 4.

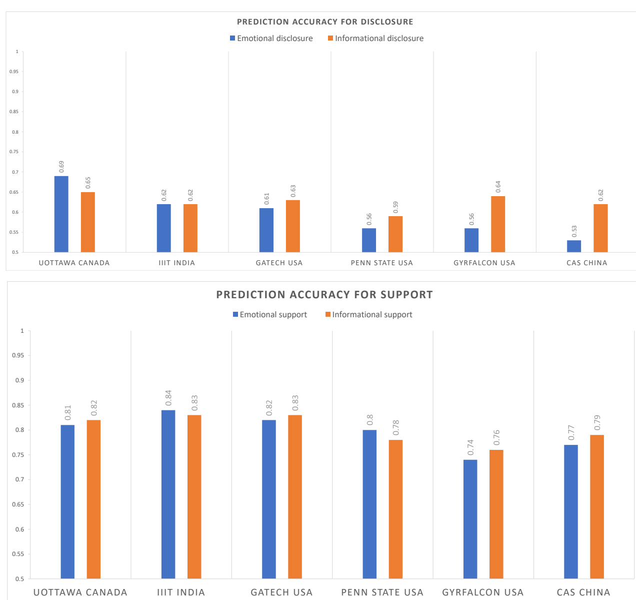

Fig. 1: Accuracy scores for the best performing system runs on Task 1 for each of the participating teams

Table 2: Systems' performance in Task 1a, ordered by their accuracy on predicting emotional disclosure.

	Emotional disclosure		Informational disclosure	
System	**Accuracy**	F_1	**Accuracy**	F_1
U.Ottawa [13] run 1	0.7	0.64	0.66	0.65
U.Ottawa [13] run 2	0.69	0.64	0.65	0.65
IIIT India run 6 [9]	0.62	0.61	0.62	0.62
GATech [4]	0.61	0.6	0.63	0.63
IIIT India run 2 [9]	0.61	0.6	0.63	0.63
IIIT India run 3 [9]	0.61	0.6	0.62	0.62
IIIT India run 4 [9]	0.61	0.6	0.62	0.62
IIIT India run 5 [9]	0.61	0.6	0.62	0.62
IIIT India run 1 [9]	0.6	0.59	0.62	0.62
IIIT India run 7 [9]	0.6	0.59	0.62	0.62
Penn State [1]	0.56	0.56	0.6	0.6
Gyrfalcon run 7 [10]	0.56	0.54	0.57	0.57
SKKU run 3 [5]	0.53	0.53	0.58	0.58
SKKU run 1 [5]	0.5	0.5	0.59	0.59
SKKU run 4 [5]	0.49	0.49	0.54	0.54
Gyrfalcon run 8 [10]	0.46	0.46	0.5	0.48
SKKU run 2 [5]	0.46	0.46	0.62	0.62
Gyrfalcon [10] run 9	0.45	0.45	0.63	0.62
Gyrfalcon [10] run 4	0.45	0.45	0.64	0.62
Gyrfalcon run 3 [10]	0.4	0.39	0.61	0.6
Gyrfalcon run 10 [10]	0.39	0.38	0.62	0.62
Gyrfalcon run 5 [10]	0.37	0.36	0.63	0.62
Gyrfalcon run 6 [10]	0.32	0.28	0.57	0.57
Gyrfalcon run 1 [10]	0.3	0.25	0.57	0.57
Gyrfalcon run 2 [10]	0.3	0.24	0.49	0.48

Four of the six systems that did Task 1 also did the bonus Task 2 to share insights based on the hidden attention or fragment layers in their deep learning models. The visualizations provided by UOttawa [13] are helpful in understanding how exactly the logical relationships between different labels are computed. Interestingly, their approach did not use any of the unlabeled data. Instead, their fragment layers appeared to infer the hierarchical relationship underlying the categories of disclosure and support. .

7 Error Analysis

We conducted a meta-analysis of system performances for Task 1 over all the sentences in the test set. When we filtered the sentences for which all or most of the approaches reported a false negative, we noted that the errors could be attributed to mislabeling, especially in the case of emotional disclosure, which had an unexpectedly high error rate. We expect that this may have happened because we transformed a 3-level annotation into a binary form; however, low-disclosure sentences may be vastly different from high-disclosure sentences. In Table 5, we provide a count of the labeling errors identified (and corrected)

Table 3: Systems' performance in Task 1b, ordered by their accuracy on predicting emotional support.

System	Emotional support Accuracy	F_1	Informational support Accuracy	F_1
IIIT run 1 [9]	0.84	0.79	0.84	0.73
IIIT run 6 [9]	0.84	0.79	0.84	0.73
IIIT run 2 [9]	0.82	0.76	0.83	0.7
IIIT run 3 [9]	0.82	0.76	0.84	0.73
IIIT run 4 [9]	0.82	0.76	0.84	0.73
IIIT run 5 [9]	0.82	0.76	0.84	0.73
IIIT run 7 [9]	0.82	0.75	0.83	0.69
GATech [4]	0.82	0.75	0.83	0.73
U.Ottawa run 2 [13]	0.81	0.75	0.82	0.73
Penn State [1]	0.8	0.72	0.78	0.48
U.Ottawa run 1 [13]	0.8	0.71	0.82	0.7
SKKU run 3 [5]	0.77	0.64	0.79	0.59
SKKU run 1 [5]	0.77	0.63	0.8	0.59
Gyrfalcon run 4 [10]	0.74	0.57	0.75	0.55
Gyrfalcon run 8 [10]	0.74	0.62	0.62	0.57
Gyrfalcon run 1 [10]	0.74	0.57	0.65	0.58
Gyrfalcon run 7 [10]	0.73	0.59	0.68	0.58
Gyrfalcon run 3 [10]	0.72	0.63	0.53	0.51
Gyrfalcon run 6 [10]	0.72	0.58	0.76	0.51
Gyrfalcon run 10 [10]	0.72	0.63	0.71	0.57
Gyrfalcon run 5 [10]	0.71	0.62	0.75	0.56
SKKU run 4 [5]	0.71	0.45	0.77	0.46
Gyrfalcon run 2 [10]	0.71	0.64	0.69	0.58
SKKU run 2 [5]	0.7	0.43	0.77	0.45
Gyrfalcon run 9 [10]	0.7	0.63	0.71	0.56

through this process. In the true spirit of a Shared Task, we have applied this feedback to identify and correct these labels. The data with corrected labels has been released. We encourage future researchers to test their approaches with the new labels.

As is expected in such tasks, other errors appeared to be because of knowledge that was implicit in a sentence and formed the basis of annotators' labels but was not directly present in the sentence. For example, "Clearly, that's disturbing for anyone to experience." was marked positive for emotional disclosure by annotators, but was predicted to be negative by most participating systems.

8 Conclusion and Future Work

The 2nd CL-Aff Shared Task AAAI-20 is the first of its kind of annotated datasets about disclosure and support in social media discussions. We have published the complete dataset to GitHub. We plan to release other labels complementary to this dataset in future tasks.

We conclude this overview with some of the main takeaways shared by our participating teams:

Table 4: Legend for Task 1 System Runs.

System No.	Run No.	Description
Gyrfalcon USA [10]	Run 1	Text only, fold 0
Gyrfalcon USA [10]	Run 2	Text only, fold 1
Gyrfalcon USA [10]	Run 3	Text only, fold 2
Gyrfalcon USA [10]	Run 4	Text only, fold 3
Gyrfalcon USA [10]	Run 5	Text only, fold 4
Gyrfalcon USA [10]	Run 6	Multimodal, fold 0
Gyrfalcon USA [10]	Run 7	Multimodal, fold 1
Gyrfalcon USA [10]	Run 8	Multimodal, fold 2
Gyrfalcon USA [10]	Run 9	Multimodal, fold 3
Gyrfalcon USA [10]	Run 10	Multimodal, fold 4
SKKU South Korea [5]	Run 1	BERT
SKKU South Korea [5]	Run 2	BERT + Emotional GLoVe
SKKU South Korea [5]	Run 3	BERT + context
SKKU South Korea [5]	Run 4	BERT + Emotional GLoVe + context
IIIT India [9]	Run 1	Model 1 (Weights to RoBERTa and ALBERT are 0 or 1)
IIIT India [9]	Run 2	Model 2 (Weights to RoBERTa and ALBERT are = .5)
IIIT India [9]	Run 3	Model 3
IIIT India [9]	Run 4	Model 4
IIIT India [9]	Run 5	Model 5
IIIT India [9]	Run 6	Finetuned RoBERTa large
IIIT India [9]	Run 7	Finetuned ALBERT xxlarge
UOttawa Canada [13]	Run 1	1024 dimensions, learning rate = 2e-5, 20 epochs
UOttawa Canada [13]	Run 2	512 dimensions, leaning rate = 2e-5, 20 epochs

Table 5: The total number of errors reported, broken down by originating forum and label

	False Positives		False Negatives	
	r/OffMyChest	**r/CasualConversation**	**r/OffMyChest**	**r/CasualConversation**
Emotional disclosure	601	371	46	20
Information disclosure	288	129	184	121
Emotional support	416	193	44	12
Information support	123	53	0	0

– UOttawa suggests that when training a model on a task using noisy datasets, it is recommended to identify and separate the data-dependent noise from the signal, and to rely on patterns and relationships based on other features. Their exemplary approach does suggest new paradigms for conceptualizing deep multi-task learning problems. However, we wonder whether the presumptions could break, for instance, when the logical relationships are accidental. In the case of GATech [4], they relied on the label distribution information to regularize their models. However, we had consciously made the decision to have a larger proportion of positive cases in the test set, which may have ultimately hurt their model performance. Perhaps the takeaway would be to look for the *semantic* relationships in the data and not rely solely on numerical trends.
– GATech reaffirms our belief in the power of semi-supervised learning for model training and prediction at scale, showing respectable performance with

an entropy-minimization approach for generated more labeled data from the unlabeled sample provided. However, they rely on the data distribution to introduce another error term to minimize the entropy of the output, and to minimize the divergence in output and input label distributions. For future modeling, we would recommend this approach only if the data generation and sampling processes are the same for both the training and the test set.
- Gyrfalcon's Super Characters approach did not appear to wholly satisfy its authors, who recommend possibly upsampling, data augmentation, or word replacement, especially when fine-tuning on small datasets.
- While it would logically be expected that adding context to models would improve model accuracy, SKKU observed no such performance gain. They recommend that rather than concatenation, adding suitable representations of context could be the right approach to enhance model performance.

Like our Shared Task last year [7], the findings do support the emerging notion about the English language as a contextualized emotional vector space, with the best performances reported by approaches that incorporated task-specific embeddings from other language models. Relying on emotional signals and the hierarchical structure of labels alone appears to have provided sufficient predictive performance. We note that in this version of the Shared Task, we did not observe any of our teams to have used syntactic information, or in building domain-specific embeddings, which were some of the more successful approaches last year.

It remains an open problem whether the models trained on this data will generalize to measure disclosure and support other platforms and conversations, and one for which we welcome future work and feedback.

Acknowledgement. Support for this research was provided by a Nanyang Presidential Postdoctoral Award and an Adobe Research Award.

References

1. Akiti, C., Rajtmajer, S., Squicciarini, A.: Contextual representation of self-disclosure and supportiveness in short text. In: Proceedings of the 3rd Workshop on Affective Content Analysis @ AAAI (AffCon2020). New York, New York (February 2020)
2. Barak, A., Gluck-Ofri, O.: Degree and reciprocity of self-disclosure in online forums. CyberPsychology & Behavior 10(3), 407–417 (2007)
3. Boyd, D.M., Ellison, N.B.: Social network sites: Definition, history, and scholarship. Journal of computer-mediated Communication 13(1), 210–230 (2007)
4. Chen, J., Wu, Y., Yang, D.: Semi-supervised models via data augmentation for classifying interactive affective responses. In: Proceedings of the 3rd Workshop on Affective Content Analysis @ AAAI (AffCon2020). New York, New York (February 2020)
5. Hyun, J., Bae, B.C., Cheong, Y.G.: [CL-Aff Shared Task] Multi-label text classification using an emotion embedding model. In: Proceedings of the 3rd Workshop on

Affective Content Analysis @ AAAI (AffCon2020). New York, New York (February 2020)
6. Jaidka, K., Guntuku, S.C., Ungar, L.H.: Facebook versus twitter: Differences in self-disclosure and trait prediction. In: Twelfth International AAAI Conference on Web and Social Media (2018)
7. Jaidka, K., Mumick, S., Chhaya, N., Ungar, L.: The CL-Aff happiness shared task: Results and key insights (2019)
8. Ma, X., Hancock, J., Naaman, M.: Anonymity, intimacy and self-disclosure in social media. In: Proceedings of the 2016 CHI conference on human factors in computing systems. pp. 3857–3869 (2016)
9. Pant, K., Dadu, T., Mamidi, R.: Bert-based ensembles for modeling disclosure and support in conversational social media text. In: Proceedings of the 3rd Workshop on Affective Content Analysis @ AAAI (AffCon2020). New York, New York (February 2020)
10. Sun, B., Yang, L., Sha, H., Lin, M.: Multi-modal sentiment analysis using super characters method on low-power cnn accelerator device. In: Proceedings of the 3rd Workshop on Affective Content Analysis @ AAAI (AffCon2020). New York, New York (February 2020)
11. Wang, Y.C., Burke, M., Kraut, R.: Modeling self-disclosure in social networking sites. In: Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing. pp. 74–85 (2016)
12. Wang, Y.C., Kraut, R., Levine, J.M.: 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. pp. 833–842 (2012)
13. Xin, W., Inkpen, D.: [CL-Aff Shared Task] Detecting disclosure and support via deep multi-task learning. In: Proceedings of the 3rd Workshop on Affective Content Analysis @ AAAI (AffCon2020). New York, New York (February 2020)
14. Yang, D., Yao, Z., Kraut, R.: Self-disclosure and channel difference in online health support groups. In: Eleventh International AAAI Conference on Web and Social Media (2017)