Identifying Differentiating Factors for Cyberbullying in Vine and Instagram

Rahat Ibn Rafiq[1], Homa Hosseinmardi[2], Richard Han[1], Qin Lv[1], and Shivakant Mishra[1]

[1] University of Colorado Boulder
{rahat.rafiq,richard.han,qin.lv,shivakaht.mishra}@colorado.edu
[2] University of Southern California, CA, USA homahoss@isi.edu

Abstract. A multitude of online social networks (OSNs) of varying types has been introduced in the past decade. Because of their enormous popularity and constant availability, the threat of cyberbullying launched via these OSNs has reached an unprecedented level. Victims of cyberbullying are now more vulnerable than ever before to the predators, perpetrators, and stalkers. In this work, we perform a detailed analysis of user postings on Vine and Instagram social networks by making use of two labeled datasets. These postings include threads of media posts and user comments that were labeled for being cyberbullying instances or not. Our analysis has revealed several important differentiating factors between cyberbullying and non-cyberbullying instances in these social networks. In particular, cyberbullying and non-cyberbullying instances differ in (i) the number of unique negative commenters, (ii) temporal distribution of positive and negative sentiment comments, and (iii) textual content of media captions and subsequent comments. The results of these analyses can be used to build highly accurate classifiers for identifying cyberbullying instances.

Keywords: cyberbullying · social computing · social network analysis

1 Introduction

The past decade has seen an unprecedented growth of Online Social Networks (OSNs). Unfortunately, this rise has also paved the way for online predators, stalkers and cyberbullying to wreak havoc on the psyche of potential victims. Cyberbullying has the potential to be more damaging than real-life bullying since it follows children and teens even outside of their schools, e.g. in their homes where they were safe earlier. The constant threat of cyberbullying in online social networks has led to devastating psychological effects in victims such as nervous breakdowns, low self-esteem, self-harm, clinical depression and in some extreme cases, suicides [5,28].

In this work, we focus on the analysis of cyberbullying on Instagram and Vine, which are especially popular with the current youth. We acknowledge that, even though Vine has been discontinued by Twitter [11], similar short video social

Fig. 1. Examples of cyberbullying in the (L) Vine and (R) Instagram online social networks.

networks are still available, such as Byte [15] and TikTok [23]. We argue that thus, the analysis of user behaviors on Vine will not be much different from the other OSNs similar to it. Cyberbullying in Instagram and Vine can happen in different ways, including sharing a humiliating/insulting/edited image/video of a victim, posting mean and hateful comments on victim's profile, including aggressive captions on shared media or hashtags, or even creating fake profiles pretending to be someone else [27]. Figure 1 provides an illustration where the profile owner is victimized by hurtful and aggressive comments posted by others in Vine and Instagram respectively. In the context of OSNs, cyber-aggression is defined as a type of behavior in an electronic context that is meant to harm another person (e.g., verbal abuse from an anonymous user online). Cyberbullying is cyber-aggression that is carried out repeatedly, against a person who cannot easily defend himself or herself, and where the bully has power over the victim [14,22]. Previous works on Instagram [7] and Vine [25] have reported that not all media sessions (shared media + associated comments) that exhibit cyber-aggression are necessarily instances of cyberbullying. In this paper, we go deeper to identify distinguishing features that differentiate cyberbullying postings from non-cyberbullying postings by conducting the following.

- Investigation of number of (i) unique commenters, (ii) unique positive sentiment commenters, and (iii) unique negative sentiment commenters
- Temporal analysis of comments belonging to the shared media
- Text-content analysis of the media captions and comments associated with the media sessions

2 Related Work

The majority of the earliest works on cyberbullying did not differentiate between instances of cyberbullying and cyber aggression [24,3,12,26,13,30,19,17,4]. This distinction is crucial since the imbalance of power in favor of the bully magnifies the effects of cyber aggression [14]. In the last few years, lots of research works have been performed in the area of efficient and accurate detection of cyberbullying using datasets that were labeled using the proper definition of cyberbullying

[32,2,8,25,29,1]. Although works have been performed to develop machine learning algorithms and features, very few works have been done to understand these features' temporal properties across a media session.

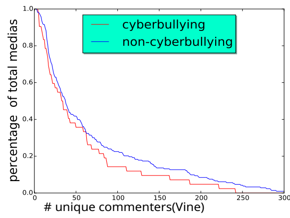
Previous research on comment analysis was mostly based on analyzing and labeling the text-content of the comments [12,26,1]. In addition to text features, the number of sent and received comments [20] and graph properties [10] were also considered to detect instances of cyberbullying. Analysis of profanity of the comments [6,4,19,17] and sentiments[31,18] in many social networks have also been explored extensively. To the best of our knowledge, none of these works explored the influence of profanity or sentiments of different parts of the comment thread (profile owner comments, media caption, etc) across a media session's temporal frame.

3 Data Set

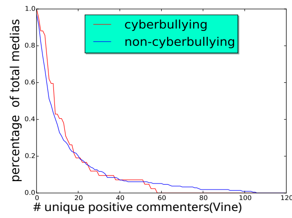
We use labeled data from Instagram [16] and Vine [25], which label each media session (shared media + associated comments) as an instance of cyberbullying, cyber-aggression, both, or neither. The data was originally collected using snowball sampling and labeled using the crowdsourcing work platform CrowdFlower (See [**?**,25] for the detailed methodology for data collection and labeling). To improve the quality of our analysis, we filter the data-set to include only media sessions with a high confidence score of being correct. For each media session, each judgment is given a trust score that incorporates the overall trust score of a labeler with the score that the labeler got while answering the test questions given on the survey (administered during the labeling process). This trust value is, in turn, incorporated with the majority voting method to assign a confidence score to the label given to a particular media session. In addition to the comment-texts and confidence score of each label, the data-sets also contain the profile owner-id, media caption, and time stamp of the shared media, timestamps of the comments, and id of the commenters belonging to the media session. For our analysis, we only use media sessions with a confidence score of 90% or higher. For Vine, this filtering reduced 983 media sessions to 42 cyberbullying media sessions and 213 non-cyberbullying media sessions. For Instagram, this filtering reduced 2216 media sessions to 239 cyberbullying media sessions and 769 non-cyberbullying media sessions. Using a high confidence score meant that the labelers were unanimous in their labeling of a particular media session.

4 Analysis of Unique Commenters

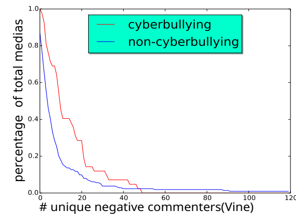
We first investigate whether the number of unique commenters has any possible influence when it comes to making a media session an instance of cyberbullying. Here, the number of unique commenters means the number of distinct users who comment on a media session. We consider the total number of unique commenters, the total number of unique positive sentiment commenters, and the total number of unique negative sentiment commenters. For this purpose, we



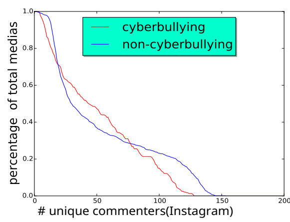
(a)



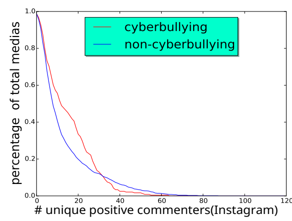
(b)



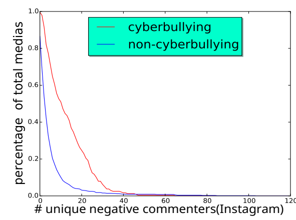
(c)



(d)



(e)



(f)

Table 1. CCDF of number of (unique,unique positive,unique negative) commenters vs percentage of total (cyberbullying,non-cyberbullying) media sessions for Vine and Instagram

take the comments associated with the labeled media sessions for both Vine and Instagram and perform sentiment analysis of all the comments using Python's NLTK library [21]. NLTK computes polarity for each comment that shows how negative or positive a particular comment's sentiment is. After getting all the comments and getting their corresponding sentiments, we generate CCDF (Complementary Cumulative Distribution Function) of the number of unique commenters (Table 1a, 1d), number of unique positive sentiment commenters (Table 1b, 1e) and the number of unique negative sentiment commenters (Table 1c, 1f) vs the percentage of total cyberbullying (non-cyberbullying) media sessions for Vine and Instagram, respectively.

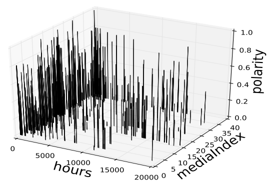
In Table 1a and 1d, the red and blue plots stand for the cyberbullying and non-cyberbullying media sessions respectively. The X-axis denotes the number of unique commenters and the Y-axis denotes the percentage of cyberbullying (non-cyberbullying) media sessions out of total cyberbullying (non-cyberbullying) media sessions having at least that many numbers of unique commenters. It is evident from the figure that for both Vine and Instagram, the number of unique commenters tends to have the same pattern for cyberbullying and non-cyberbullying. The same indistinguishable trend for both labels is also seen for the total number of unique positive sentiment commenters from Table 1b and 1e. This means that *for both Vine and Instagram, cyberbullying, and non-cyberbullying media sessions tend to have the same trend when it comes to the number of unique commenters and the number of unique positive sentiment commenters.*

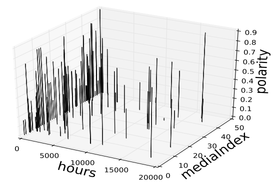
However, for the number of unique negative commenters (Table 1c and 1f), cyberbullying and non-cyberbullying sessions differ from one another. It is seen that, for both Vine and Instagram, the number of unique negative commenters trend for cyberbullying media sessions fall much more slowly than for non-cyberbullying sessions. The figure shows that the percentage of cyberbullying media sessions that have at least a certain number of negative unique commenters is much more than that of non-cyberbullying media sessions. This means that *cyberbullying media sessions are likely to have more unique negative sentiment commenters for Vine and Instagram*. We believe this is because, in a cyberbullying media session, perpetrators often gang up against the victim and thus spikes up the number of unique negative sentiment commenters. It can also be seen that after 40 unique negative sentiment commenters, the non-cyberbullying trend starts to show a long tail, which is not seen for the cyberbullying trend. This is because some non-cyberbullying media sessions belong to celebrities and famous brands that have a large number of comments from a large number of followers, and sometimes the commenters express awe with expletives and/or swear words in those media sessions, thus contributing to the long tail.

5 Temporal Analysis of Negative and Positive Sentiment Comments

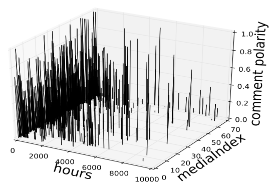
Now we turn our attention to the temporal analysis of comments on a particular media session since the media session is shared. We perform the analysis on all negative and positive sentiment comments where the sentiment was determined by using Python's NLTK library[21]. We do the temporal analysis for both negative and positive sentiment comments because we think media sessions that are tagged as cyberbullying are more likely to have a higher concentration of negative sentiment comments and a lower concentration of positive comments, thus resulting in the imbalance of power as per the definition of cyberbullying.

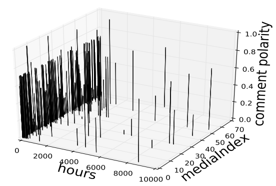
Figures in table 2 show the temporal comment polarity for all negative sentiment comments for a particular cyberbullying(not cyberbullying) media session since the sharing of the media session for Vine and Instagram respectively. It is evident from the figures that the negative sentiment comments are much more spread up across the temporal frame of each media session in the case of cyberbullying sessions than for the non-cyberbullying sessions. The cyberbullying media sessions have a constant flow of high negative sentiment comments pouring in, even after a considerable amount of time since the sharing of the media. On the contrary, the same cannot be said for the non-cyberbullying sessions as the number of negative sentiment comments tend to go down as time moves on. We believe this is a very important factor that can differentiate a cyberbullying media session from a non-cyberbullying one. This shows that *in the cyberbullying media sessions, the negative sentiment comments persist even after a long time since the sharing of the media, which confirms the factor of repetition of aggression in the definition of cyberbullying.*





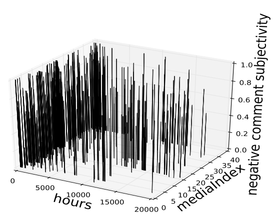
(a) vine bullying polarity (b) vine notbullying polarity

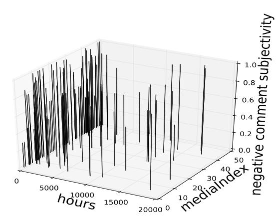




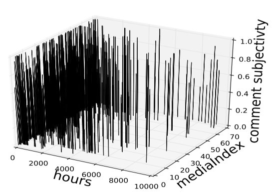
(c) inst. bullying polarity (d) inst. notbullying polarity

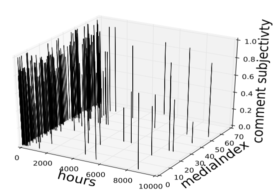
Table 2. Polarity of negative sentiment comments as time moves on since the media session has been posted for cyberbullying and non-cyberbullying media sessions in Vine and Instagram





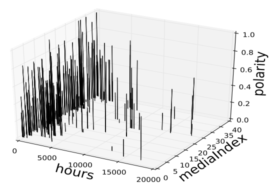
(a) vine bullying subjectivity (b) vine notbullying subjectivity



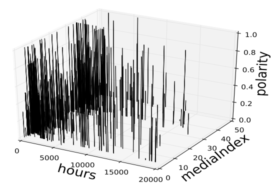


(c) inst. bullying subjectivity (d) inst. notbullying subjectivity

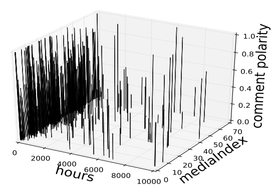
Table 3. Subjectivity of negative sentiment comments as time moves on since the media session has been posted for cyberbullying and non-cyberbullying media sessions in Vine and Instagram



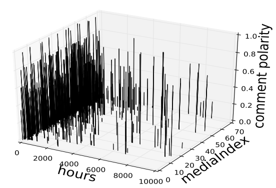
(a) vine bullying polarity



(b) vine notbullying polarity



(c) inst. bullying polarity



(d) inst. notbullying polarity

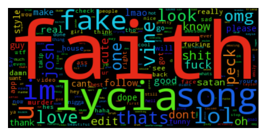
Table 4. Polarity of positive sentiment comments as time moves on since the media session has been posted for cyberbullying and non-cyberbullying media sessions in Vine and Instagram

Next, we conduct the same kind of temporal analysis to investigate the subjectivity of the negative sentiment comments for cyberbullying and non-cyberbullying media sessions for both Vine and Instagram. Subjectivity determines how severe a negative sentiment comment is [21]. The intuition is that the cyberbullying media sessions should have more negative sentiment comments with comparatively higher subjectivity, thus being more aggressive which in turn results in cyberbullying. Figures in table 3 show the subjectivity values of all the negative sentiment comments posted for the cyberbullying and non-cyberbullying media sessions since the sharing of the media sessions for both Vine and Instagram. It is apparent from the figures that the cyberbullying media sessions for both Vine and Instagram keep having negative sentiment comments with very high subjectivity spread across the temporal frame since the sharing of the media session. This results in the denser concentration of high bars for the cyberbullying sessions. *So not only the cyberbullying media sessions keep getting more negative sentiment comments even after a long time since the media session is posted, but also the negative sentiment comments tend to have more subjectivity than non-cyberbullying media sessions.*

Now, we conduct a temporal analysis of the polarity of all the positive sentiment comments for cyberbullying and non-cyberbullying media sessions for both Vine and Instagram. The expectation is that the cyberbullying sessions should have a less concentrated positive sentiment comments, thus rendering the effect of the imbalance of power as delineated in the definition of cyberbullying. Figures in table 4 show the temporal comment polarity for all positive sentiment comments for a particular media session since the moment the media session has been posted for Vine and Instagram respectively. From the figures, it is seen that the *density of positive comments coming in for cyberbullying media sessions for both Vine and Instagram is much less than the non-cyberbullying media sessions*. This lesser concentration of positive sentiment comments coupled with the denser concentration of negative sentiment comments with high subjectivity spread across the temporal frame instigates the effect of the imbalance of power and repeated aggression, thus rendering the media session a cyberbullying one.

6 Analysis of Comments

(a) vine bullying words



(b) vine not bullying words

(c) insta. bullying words

(d) insta. notbullying words

(e) vine bullying idf

(f) vine notbullying idf

(g) insta. bullying idf

(h) insta. notbullying idf

Table 5. Frequency distribution and top idf valued distribution of words used for the cyberbullying and non-cyberbullying media sessions' comments in Vine and Instagram.

Next, we perform a text-content analysis for the comments associated with a media session for both Vine and Instagram. We consider the comments associated with a media session from part of a discussion thread, and our goal is to determine the differences between a discussion thread of a cyberbullying session and a discussion thread of a non-cyberbullying thread.

First, we devise a word frequency cloud for the cyberbullying and non-cyberbullying media sessions for both the social networks to get an idea of the words that occur frequently. Figures in Table 5a, 5b, 5c, 5d show the frequency distribution of words of all the media sessions' comments belonging to cyberbullying and non-cyberbullying media sessions for Vine and Instagram respectively. It can be seen from these figures that *negative sentiment words are much more frequent in the discussion comment threads of cyberbullying sessions*.

Next, we do an IDF (Inverse Document Frequency) analysis of the media sessions' comments that measures how common a word is across all media session comment discussions for cyberbullying and non-cyberbullying sessions. The difference between the frequency analysis and IDF analysis is that frequency analysis only takes into account the number of times a word appears in a discussion thread whereas IDF analysis gives us words that are common across all cyberbullying and non-cyberbullying comment discussion threads. Thus, a word that appears 10 times in 10 different documents will have lower IDF than a word that appears 10 times in a single document. Figures in Table 5e, 5f, 5g and 5h show the commonly appearing words for cyberbullying and non-cyberbullying media session comment threads for both Vine and Instagram respectively across all the corresponding media session comment threads. The bigger a word is in the word cloud, the more common it is across all the media session comment threads belonging to either cyberbullying or non-cyberbullying label. It is evident that, as it was seen also from the previous paragraph, *a cyberbullying media session comment discussion thread is much more likely to have negative sentiment words.*

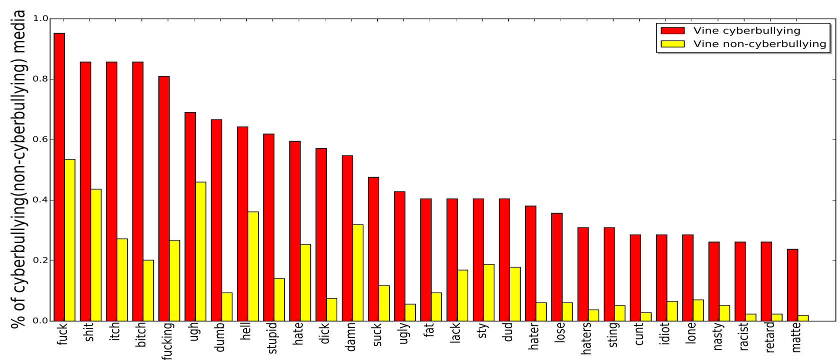


Fig. 2. Negative sentiment words vs percentage of cyberbullying (non-cyberbullying) media sessions out of total cyberbullying (non-cyberbullying) media sessions' comment threads containing that word in Vine.

To further confirm the aforementioned claim, we use the negative sentiment word list [9] and find out the percentage of cyberbullying (non-cyberbullying) media sessions out of total cyberbullying (non-cyberbullying) media sessions

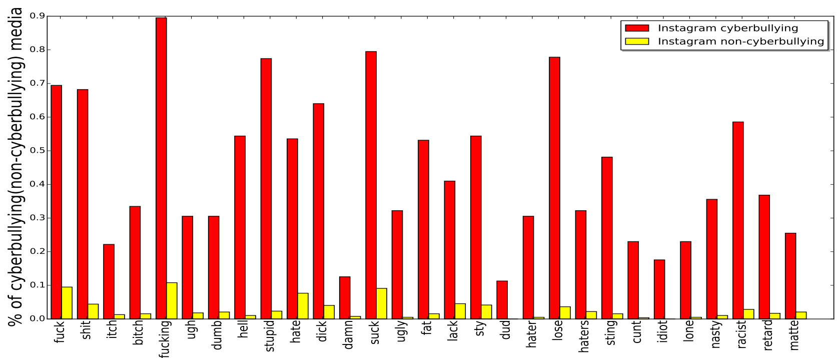


Fig. 3. Negative sentiment words vs percentage of cyberbullying (non-cyberbullying) media sessions out of total cyberbullying (non-cyberbullying) media sessions' comment threads containing that word in Instagram.

whose comment threads contain those negative sentiment words. We intuit that negative sentiment words appear more in the cyberbullying media sessions than the non-cyberbullying media sessions for both the social networks, thus forming a differentiating factor for cyberbullying. We can see from Figures 2 and 3, negative sentiment words are much more likely to appear in a cyberbullying media session's associated comments than the non-cyberbullying media sessions, thus further confirming our claim:*a cyberbullying media session comment discussion thread is much more likely to have negative sentiment words*.

7 Analysis of Media Captions

In this section, we analyze the captions that the profile owners put for each media-sessions when they share the media in Vine and Instagram. We intuit that these media captions set the topic of the discussion that comes after the media is shared and thus setting precedent for the oncoming comment threads. We do this analysis to check if there are any differentiating topics, words, or subjects when it comes to cyberbullying and not-cyberbullying media sessions for both Vine and Instagram.

For this analysis, we use the negative sentiment word list [9] and find out the percentage of cyberbullying and not-cyberbullying media sessions out of total cyberbullying and not-cyberbullying media sessions respectively whose captions contain those negative sentiment words. Our intuition is that negative sentiment words appear more in the cyberbullying media sessions' captions than the not-cyberbullying media sessions' captions, thus forming a differentiating factor for cyberbullying. Surely enough, as it can be seen from Figure 4, negative sentiment words are much more likely to appear in a cyberbullying media session's associated caption than the not-cyberbullying media session's captions, thereby further supporting our claim: *a cyberbullying media session caption is much more likely to have negative sentiment words*.

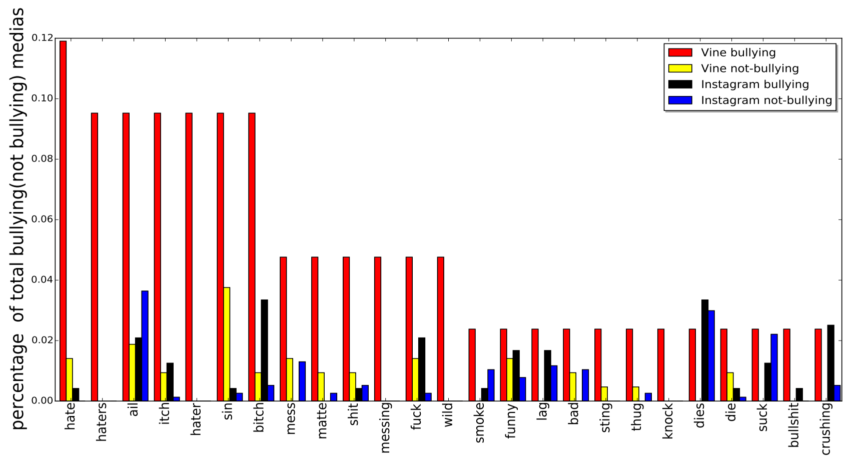


Fig. 4. Negative sentiment words vs percentage of bullying(not bullying) media sessions out of total cyberbullying (not- cyberbullying) media sessions' captions containing that word in Vine and Instagram

8 Conclusions

To the best of our knowledge, this is the first paper to investigate factors that differentiate a cyberbullying session from a non-cyberbullying one for both Vine and Instagram, two media-based online social networks leveraging labeled data that used appropriate definition of cyberbullying. We analyze the number of unique commenters, unique positive sentiment commenters, and unique negative sentiment commenters. We then perform a temporal analysis of all comments for both social networks. Finally, we conduct a content analysis of the comment threads and media-captions belonging to the labeled cyberbullying and non-cyberbullying media sessions.

The key findings of this research are as follows. First, for both Vine and Instagram, cyberbullying media sessions are more likely to have more unique negative sentiment commenters. Second, in the cyberbullying media sessions, negative sentiment comments persist with higher subjectivity even after a long time since the media has been posted, which is not the case for non-cyberbullying media sessions. Third, the density of positive comments coming in for cyberbullying media sessions for both Vine and Instagram is much less than that for the non-cyberbullying media sessions across the temporal frame. Fourth, the comment discussion threads across time units belonging to cyberbullying media sessions show a high level of negative sentiment polarity than those belonging to non-cyberbullying sessions. Fifth, while for non-cyberbullying media sessions, negative sentiment discussions tend to fizzle out as time moves on, that is not the case for cyberbullying media sessions in Vine and Instagram. Sixth, a cyberbullying media session's media caption is more likely to have negative sentiment words. Seventh, a cyberbullying media session comment thread is much more likely to have negative sentiment words than a non-cyberbullying media session.

In future, we plan to leverage these insights to build a highly accurate cyberbullying classifier for both Vine and Instagram.

9 Acknowledgements

This work was supported by the US National Science Foundation (NSF) through grant CNS 1528138.

References

1. Arslan, P., Corazza, M., Cabrio, E., Villata, S.: Overwhelmed by negative emotions? maybe you are being cyber-bullied! In: Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing. p. 1061–1063. SAC '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3297280.3297573, https://doi.org/10.1145/3297280.3297573
2. Cheng, L., Li, J., Silva, Y.N., Hall, D.L., Liu, H.: Xbully: Cyberbullying detection within a multi-modal context. In: Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining. p. 339–347. WSDM '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3289600.3291037, https://doi.org/10.1145/3289600.3291037
3. Dadvar, M., de Jong, F.M.G., Ordelman, R.J.F., Trieschnigg, R.B.: Improved cyberbullying detection using gender information. In: Twelfth Dutch-Belgian Information Retrieval Workshop, DIR. pp. 23–25. University of Ghent (2012)
4. Dinakar, K., Reichart, R., Lieberman, H.: Modeling the detection of textual cyberbullying. (2011)
5. Goldman, R.: Teens indicted after allegedly taunting girl who hanged herself, bbc news. http://abcnews.go.com/Technology/TheLaw/teens-charged-bullying-mass-girl-kill/story?id=10231357 (2010), [Online; accessed 14-January-2014]
6. Hosseinmardi, H., Ghasemianlangroodi, A., Han, R., Lv, Q., Mishra, S.: Towards understanding cyberbullying behavior in a semi-anonymous social network. In: IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). pp. 244 – 252. IEEE, Beijing,China (2014)
7. Hosseinmardi, H., Rafiq, R.I., Han, R., Lv, Q., Mishra, S.: Prediction of cyberbullying incidents in a media-based social network. In: Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. IEEE, San Francisco,CA,USA (2016)
8. Hosseinmardi, H., Rafiq, R.I., Han, R., Lv, Q., Mishra, S.: Prediction of cyberbullying incidents in a media-based social network. In: Advances in Social Networks Analysis and Mining (ASONAM), 2016 IEEE/ACM International Conference on. pp. 186–192. IEEE (2016)
9. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 168–177. ACM (2004)

10. Huang, Q., Singh, V.K., Atrey, P.K.: Cyber bullying detection using social and textual analysis. In: Proceedings of the 3rd International Workshop on Socially-Aware Multimedia. pp. 3–6. ACM, New York, NY, USA (2014). https://doi.org/10.1145/2661126.2661133, http://doi.acm.org/10.1145/2661126.2661133
11. Jr, T.H.: Twitter is officially shutting down vine today. https://fortune.com/2017/01/17/twitter-shut-down-vine-tuesday/ (2017), [Online; accessed July 28, 2020.]
12. K. Reynolds, A.K., Edwards, L.: Using machine learning to detect cyberbullying. Proceedings of the 2011 10th International Conference on Machine Learning and Applications and Workshops **2**, 241–244 (2011). https://doi.org/http://doi.ieeecomputersociety.org/10.1109/ICMLA.2011.152
13. Kontostathis, A., Reynolds, K., Garron, A., Edwards, L.: Detecting cyberbullying: query terms and techniques. In: Proceedings of the 5th Annual ACM Web Science Conference. pp. 195–204. ACM (2013)
14. Kowalski, R.M., Limber, S., Limber, S.P., Agatston, P.W.: Cyberbullying: Bullying in the digital age. John Wiley & Sons, Reading, MA. (2012)
15. Law, T.: Vine has a new successor: The 6-second video app byte. https://time.com/5771854/vine-byte-app-launch/ (2020), [Online; accessed July 28, 2020.]
16. Li, H.H.S., Yang, Z., Lv, Q., Han, R.I.R.R., Mishra, S.: A comparison of common users across instagram and ask.fm to better understand cyberbullying. In: 2014 IEEE Fourth International Conference on Big Data and Cloud Computing. pp. 355–362. IEEE, Sydney, Australia (Dec 2014). https://doi.org/10.1109/BDCloud.2014.87
17. Nahar, V., Li, X., Pang, C.: An effective approach for cyberbullying detection (2013)
18. Nahar, V., Unankard, S., Li, X., Pang, C.: Sentiment analysis for effective detection of cyber bullying. In: Web Technologies and Applications, pp. 767–774. Springer (2012)
19. Nahar, V., Unankard, S., Li, X., Pang, C.: Semi-supervised learning for cyberbullying detection in social networks. In: Databases Theory and Applications. pp. 160–171. LNCS'12 (2014)
20. Nalini, K., Sheela, L.J.: Classification of tweets using text classifier to detect cyber bullying. In: Emerging ICT for Bridging the Future-Proceedings of the 49th Annual Convention of the Computer Society of India CSI Volume 2. pp. 637–645. Springer (2015)
21. NLTK: Python nltk library. https://www.nltk.org/ (2020), [Online;accessed March 30, 2020]
22. Patchin, J.W., Hinduja, S.: An update and synthesis of the research. Cyberbullying Prevention and Response: Expert Perspectives p. 13 (2012)
23. Perez, S.: It's time to pay serious attention to tiktok. https://techcrunch.com/2019/01/29/its-time-to-pay-serious-attention-to-tiktok/ (2019), [Online; accessed July 28, 2020.]
24. Ptaszynski, M., Dybala, P., Matsuba, T., Masui, F., Rzepka, R., Araki, K., Momouchi, Y.: In the service of online order tackling cyberbullying with machine learning and affect analysis. International Journal of Computational Linguistics Research **1**(3), 135–154 (2010)
25. Rafiq, R.I., Hosseinmardi, H., Han, R., Lv, Q., Mishra, S., Mattson, S.A.: Careful what you share in six seconds: detecting cyberbullying instances in vine. In: Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015. pp. 617–622. ACM, Paris,France (2015)

26. Sanchez, H., Kumar, S.: Twitter bullying detection. In: NSDI. pp. 15–15. USENIX Association, Berkeley, CA, USA (2012)
27. Silva, T.H., de Melo, P.O.S.V., Almeida, J.M., Salles, J., Loureiro, A.A.F.: A picture of Instagram is worth more than a thousand words: Workload characterization and application. In: 2013 IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS). pp. 123–132. IEEE (2013)
28. Smith-Spark, L.: Hanna smith suicide fuels calls for action on ask.fm cyberbullying, cnn. `http://www.cnn.com/2013/08/07/world/europe/uk-social-media-bullying/` (2013), [Online;accessed 14-January-2014]
29. Soni, D., Singh, V.K.: See no evil, hear no evil: Audio-visual-textual cyberbullying detection. Proc. ACM Hum.-Comput. Interact. **2**(CSCW) (Nov 2018). https://doi.org/10.1145/3274433, `https://doi.org/10.1145/3274433`
30. Xu, J.M., Jun, K.S., Zhu, X., Bellmore, A.: Learning from bullying traces in social media. In: NAACL HLT. pp. 656–666. Association for Computational Linguistics (2012)
31. Xu, J.M., Jun, K.S., Zhu, X., Bellmore, A.: Learning from bullying traces in social media. In: Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. p. 656–666. NAACL HLT '12, Association for Computational Linguistics, USA (2012)
32. Yao, M., Chelmis, C., Zois, D.: Cyberbullying ends here: Towards robust detection of cyberbullying in social media. In: The World Wide Web Conference. p. 3427–3433. WWW '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3308558.3313462, `https://doi.org/10.1145/3308558.3313462`