

Prospectus

Detection of Inappropriate Social Media Posts Using Emotion Detection
(Technical Topic)

Prevalence of Cyberbullying in Social Media and its Effect on Youth
(STS Topic)

By

Rachel Lew

May 4, 2021

Technical Project Team Members: N/A

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: _____ *Rachel Lew* _____

Technical Advisor: ~~Nada Basit~~ *Nada Basit*

STS Advisor: _____ *E. D. Wylie* _____

Introduction

As technology has advanced, so have the ways in which we interact with one another. Communicating online through social media platforms has become more common and normalized as an everyday occurrence, especially amongst young adults and even children. Screen use in adolescents continues to increase while some alarming statistics seem to follow the same trajectory, like the “rates of suicide among youth aged 10-24 increasing 56% from 2007 to 2017” (Nesi, 2020, p. 116). This statistic is not as alarming knowing that children have been introduced to the toxic side of social media at an early age, including cyberbullying and online harassment. In fact, studies have witnessed that such exposures to online violence have negative psychological effects on children (Patton, 2014).

So, how is this aftermath handled or controlled? While we would think to deem the parents responsible for teaching their children the appropriate way to use social media, there is actually a large disconnect between parents and their youth regarding social media, as parents “may lack a basic understanding of these new forms of socialization” that “are integral to their children’s lives” (O’Keeffe, 2011, p. 801). This then prompts the question of, who is really responsible for the consequences related to social media use? On top of that, little is known about how social media platforms in general monitor cyberbullying among their young users, or whether this monitoring even exists. Thus, I will be exploring how to mitigate and regulate cyberbullying and its adverse effects on youth from social media platforms by engaging with peer-reviewed papers. Through my technical topic, I will explore a solution to mitigate cyberbullying before it even begins, and through my STS topic, I will explore a solution to tackle the effects cyberbullying has on youth.

Technical Topic

I believe it is important for social media to be monitored for any signs of cyberbullying because of the very real effects it can cause on children. When is it the point where online threats and behaviors become real instances to be concerned of? Just because trauma is occurring online does not mean there will not be effects when the individual is offline. Without being able to flag such online behaviors, people cannot be warned about the detrimental effects that could occur real time, such as extreme depression, school shootings, and suicide. So, how can technology be used to help minimize and regulate cyberbullying to prevent real world repercussions? I propose to use technology for emotion detection while users are engaging in social media activity. Then, I will use such data to flag certain posts or comments the user is simultaneously viewing while feeling negative emotions.

As cyberbullying has affected more and more adolescents, the detection of such posts and comments have become more prominent in research. Van Bruwaene et al. (2020) use machine learning and a text-based method to detect cyber aggression and bullying posts. While their study had significant improvements in the accuracy of detecting such posts, there is still much more to research, since adjusting certain factors drops the performance of the algorithms. Currently, the main way for cyber detection seems to be through text-based methods, where researchers expand this method by also including other features, like user-based features (Van Bruwaene et al., 2020, p. 853). Thus, the expansion of emotion detection has not yet been investigated within the topic of cyber detection, which I am planning to explore.

Engineers have already been researching ways to detect emotion through smart devices by using physiological signals (Kanjo et al., 2015). Kanjo et al. (2015) suggest that emotions are triggered and can be sensed through data like heart rate, skin temperature, and blood pressure. Studies such as Yamamoto et al. (2009) also use physiological signals in respect with a feeling distinction model to compare how an individual is reacting with a certain emotion. Because such methods have been researched and tested, I am hoping to apply this methodology to also include a simultaneous tracking system of an individual's social media activity. Because an individual's emotions can tell us how a post may be received, I am hoping that by learning an individual's emotional response to a post can help determine whether the post should be singled out. By having this dual system, I can create an approach to flag or create concern surrounding posts on social media that may be creating a negative effect on an individual.

STS Topic

As social media has become a daily recreation for many young adults, there is much to be learned about how their interactions with each other online affect their mental well-being. Nesi, Choukas-Bradley, and Prinstein (2018) argue the significance of certain features of social media as factors in the changing relations among youth. Some important features they focus on include asynchronicity, permanence, publicness, availability, cue absence, quantifiability, and visualness (Nesi et al., 2018, p. 275). These features were selected because they have higher levels of intensity in social media compared to in person interactions. Because social media has created differences in interactions that can affect the mindset of youth, it is important to study these negative effects because of the changing patterns in how adolescents view relationships in accordance with social media.

Interestingly, Nesi et al. (2018) propose a theoretical framework called a transformative framework to understand the effects social media has on youth and their relationships. With this framework, they make a good point where there are a lot of trivial behaviors on social media that can be seen as “bullying” in the eyes of the victim. For example, they point out how simple acts such as liking a post can show support (Nesi et al., 2018). An individual can construe someone not liking their post as mean or hurtful. With this in mind, how much can social media platforms truly regulate? Because of all of this unintentional damaging behavior that may be going on, is there really much that the engineers who developed the platform can do to minimize the negative effects on an individual’s mental health? Even Underwood and Ehrenreich (2017) emphasize how the consequences of cyber aggression do not have to be from repetitive events of cyberbullying, but that even one instance is enough to cause traumatic effects on a victim.

Thus, detection of trauma caused by social media can seem minor, making it even more difficult to detect for social media platforms, in which the only way to possibly find such small instances would be through manual monitoring. Monitoring cyberbullying and negative effects may be too difficult to do for social media platforms in a large scale, so is this responsibility for mitigating cyberbullying then turned to parents, school officials, and other children specialists? This would rely on victims being able to open up to the adult figures in their life. However, Hoff and Mitchell (2009) also found that students are reluctant to go to school officials or their parents about the cyberbullying because of fear, embarrassment, and the stigma that the officials do not care or will not help. This reluctance highlights how there is a gap between adults and students when it comes to social media and understanding the full effects of it.

By exploring the ways in which social media can impact an individual, I hope to determine how much can be done to mitigate these negative effects on mental health. Whether it be flagging offensive posts or having to turn to an individual's adult figures for aid, a solution for lessening the adverse effects of social media due to cyberbullying, no matter how minor the incident, is extremely important for the overall mental well-being of our youth.

Conclusion

To summarize, I hope to find a solution to help mitigate cyberbullying on social media through my technical deliverable by creating a device that can detect a change in one's emotions to flag what they are viewing on the social media platform. I hope to use my STS deliverable to determine how the negative impact on youth's mental health can be mitigated, or if it can be at all. Overall, my goal is to bring more awareness to social media and its effect on youth. Specifically, I hope to show how easily bullying can be translated into an online environment, that it can exist in both explicit and implicit forms, and thus causing an effect on one's mental health.

References

- Abi-Jaoude, E., Naylor, K. T., & Pignatiello, A. (2020). Smartphones, social media use and youth mental health. *Cmaj*, 192(6), E136-E141.
- Clerkin, E. M., Smith, A. R., & Hames, J. L. (2013). The interpersonal effects of Facebook reassurance seeking. *Journal of affective disorders*, 151(2), 525-530.
- Davila J, Hershenberg R, Feinstein BA, Gorman K, Bhatia V, & Starr LR (2012, April 1). Frequency and Quality of Social Networking Among Young Adults: Associations With Depressive Symptoms, Rumination, and Corumination. *Psychology of Popular Media Culture*, 1(2), 72 - 71. Retrieved from <https://doi.org/10.1037/a0027512>
- Dianne L. Hoff, & Sidney N. Mitchell (2009, August 14). Cyberbullying: causes, effects, and remedies. *Journal of Educational Administration*, 47(5), 652 - 665.
- Kanjo, E., Al-Husain, L., & Chamberlain, A. (2015). Emotions in context: examining pervasive affective sensing systems, applications, and analyses. *Personal and Ubiquitous Computing*, 19(7), 1197-1212.
- Matsubara, M., Augereau, O., Sanches, C. L., & Kise, K. (2016, December). Emotional arousal estimation while reading comics based on physiological signal analysis. In *Proceedings of the 1st International Workshop on coMics ANalysis, Processing and Understanding* (p. 7). ACM.
- Mitchell, K. J., & Ybarra, M. L. (2007). Online behavior of youth who engage in self-harm provides clues for preventive intervention. *Preventive Medicine*, 45(5), 392-396. <https://doi.org/10.1016/j.ypmed.2007.05.008>
- Nesi, J. (2020). The impact of social media on youth mental health: challenges and opportunities. *North Carolina medical journal*, 81(2), 116-121.
- Nesi, J., Choukas-Bradley, S., & Prinstein, M. J. (2018). Transformation of Adolescent Peer Relations in the Social Media Context: Part 1—A Theoretical Framework and Application to Dyadic Peer Relationships. *Clinical Child and Family Psychology Review*, 21, 267-294.
- O'Keeffe, G. S., & Clarke-Pearson, K. (2011). The impact of social media on children, adolescents, and families. *Pediatrics*, 127(4), 800-804.
- Patton, D. U., Hong, J. S., Ranney, M., Patel, S., Kelley, C., Eschmann, R., & Washington, T. (2014). Social media as a vector for youth violence: A review of the literature. *Computers in Human Behavior*, 35, 548-553.
- Sherman, L. E., Payton, A. A., Hernandez, L. M., Greenfield, P. M., & Dapretto, M. (2016). The Power of the Like in Adolescence: Effects of Peer Influence on Neural and Behavioral Responses to Social Media. *Psychological Science*, 27(7), 1027–1035. <https://doi.org/10.1177/0956797616645673>
- Underwood, M. K., & Ehrenreich, S. E. (2017). The Power and the Pain of Adolescents' Digital Communication: Cyber Victimization and the Perils of Lurking. *American Psychologist*, 72(2), 144-158. <https://doi.org/10.1037/a0040429>
- Van Bruwaene, D., Huang, Q., & Inkpen, D. (2020). A multi-platform dataset for detecting cyberbullying in social media. *Language Resources and Evaluation*, 54(4), 851-874.
- Yamamoto, J., Kawazoe, M., Nakazawa, J., Takashio, K., & Tokuda, H. (2009). MOLMOD: analysis of feelings based on vital information for mood acquisition. *Skin*, 4(6).