**ABSTRACT**

Social networking sites are primarily used for communicating and connecting with other people. However, some people use these technologies to harm others emotionally. Cyberbullying refers to the use of technology to humiliate and slander behaviour towards other people with the use of computers or mobile phones. It takes form of hate messages sent through social media and e-mails. Truly, cyberbullying has become rampant in many countries. In order to address this issue, the team proposed a method to detect cyberbullying entries on social media through the use of Natural Language Processing (NLP). The data that was used by the researchers was collected from Youtube, Twitter, and Facebook. Each statement that was extracted from these social networking sites were ranked according to their harmfulness level (wherein 0, indicates no indication of cyberbullying event, 1 indicates mild cyberbullying occurrences and 2 indicates severe cyberbullying occurrences). Furthermore, it was annotated into six cyberbullying categories: bad description, intelligence, physical appearance, race and culture, sexuality, and social rejection. Among the 450 words that were extracted from 600 statements, bad description was most prevalent with a frequency of 27%.

Keywords: *Cyberbullying, Detection, Implications, Social Media*