NLP Project: Paperwork

**INTRODUCTION**

* 1. **Background of the Problem**

Long before men evolved into species of higher intellectual capabilities, bullying was believed to have been evident. Boehm (2012) stated in his book, *Moral Origins*, that primates, specifically monkeys and chimpanzees, frequently execute bullying-like deportment against members of their own kind, which would therefore provide them an edge in terms of social stature, acquired resources, and reproductive "opportunities" among the rest. Upon the rise of the *Homo-sapiens* (the genus into which humans of today are classified), the purpose of bullying was redefined from social dominance into a mere destructive act. Hogan Sherrow, an anthropologist, believes that "the ability of language to facilitate communications, coordinate behaviors, express thoughts and gossip has completely altered the intensity and impact of bullying". Fast-forward to the 21st century, likewise known as the era of widespread technological advancements, a new form of bullying emerges - cyberbullying. Cyberbullying, also referred to as "modern-day bullying" (Marusarz & Tienabeso, 2010), remains as a serious issue particularly among the adolescents of this era. In a study conducted by the students of Paris West University Nanterre La Défense on a sample of 272 adolescents, 39.4% of them were determined to have been victims of the said type of bullying (with the other 16.9% as cyberbullies). Similarly, an author affiliated with the Department of Psychology in Suffolk University, Boston, MA claimed that the emergence of cyberbullying as a powerful tool for causing harm among teenagers and their peers is due to the latter's frequent use of social media. (Horner, Asher & Fireman, 2015, pp. 288-295).

With the immense number of new gadgets being introduced into the market almost every year and the accessibility of acquiring a reliable internet connection, the probability of people engaging in different social media websites, forums, blogs or other forms of social communities online are not likely to decrease. Consequently, it leads to the formation of virtual “hang-outs” of some sort. And whenever groups of people are involved, specifically in areas where adult supervision is limited (Will & Clayburn, 2010), the occurrence of cyberbullying becomes inevitable. The alarming fact about cyberbullying is that it can be done by anyone (including people whom the victim is not familiar with), in an instant, may spread across different areas, and may harm a person without other people’s knowledge. (Stopbullying.gov, 2014)

From being dubbed as the “Texting Capital of the World” to “Social Media Capital”, the Philippines had proven itself enough to be recognized as an overly social country. (Hansen, 2010). As of 2015, there are about 47 000 000 Filipino users who are active on Facebook according to InternetWorldStats.com (2015). This particular norm of owning at least one personal account in any of the renowned social media sites continues to expand day by day. While the existence of these particular types of media provided ample benefits with regard to improving former communication-related processes, such sites have likewise been considered as the launch-point of common cyberbullying assaults occurring within the country.

Oddly enough, despite the dangers cyberbullying can inflict on an individual, only a small number of reports are continuously being submitted to designated authorities. Dr. Ryan Guinaran, Ph.D. claimed that the latter was due to the fact that cyberbullying in the Philippines (in comparison to other countries) tends to be more on a conservative level (Witkus, 2012). If Filipinos continue to practice this type of passive attitude regarding the matter at hand, then even with the efforts granted by the government and NGOs alike, cyberbullying will still persist. Thus, instead of waiting for the parties involved to voluntarily explain their side to the people concerned, the group had the thought of taking advantage of the same platform where the aforementioned event was known to have been rampant – technology – as its own countermeasure.

* 1. **Statement of the Problem**

How can Natural Language Processing be used to detect trends of cyberbullying in social media sites among Filipino adolescents?

**1.3 Objectives**

The main objective that the group is trying to achieve in this project is to be able to determine how the Natural Language Processing (NLP) technology can be used to aid in the efforts of people in preventing the occurrence of cyberbullying to intensify within social media networking sites among adolescents by detecting its trends. The results from the said detection should be reflected toward the user with the help of ‘flags’.

**1.4 Significance**

In the Philippines, nowadays, the younger population tend to visit social media websites more frequently than the rest (Get Hooked 360, 2013). That fact alone is upsetting. Adolescents have a tendency to deal with things impetuously due to their immaturity. They are most likely unable to identify the intensity of the damage that they had done until it finally occurred (KidsHealth, 2016).

At the same time, several studies indicated that most adolescents are reluctant to admit to being the victims of cyberbullying. Each of them have their own specific reasons for not doing so, making the situation even worse (KidsHealth, 2016). In fact, to fight the effects of cyberbullying, one has to have a mental well-being. People are capable of erecting a strong façade, however it is a lot difficult than it seemed. (Kreger, 2015). If a person has already reached his or her limit, the latter will be more probable to search for ways of escapism (e.g. the use of alcohol and drugs, skipping school, suicide, picking-on weaker peers, etc.), leading him/her to the path of unproductive growth. Additionally, this particular effect can spread among people whom the victim frequently interacts with. (Hinduja & Patchin, 2008)

**1.5 Scope and Limitations**

This project will focus on detecting and analyzing cyberbullying occurrences, which will revolve among issues deemed sensitive amongst the Filipino population, in public Facebook “posts” exchanged between adolescents (people whose age falls between 10-19 years old) living in Metro Manila, Philippines. There are 2 main actors involved alongside the automatic detection process – adolescents, which may either be the bully, the bullied, or the bystander, and the moderators, which are preferably adults (e.g. parents of the teens involved) or any person who was authorized to monitor such events in a particular social media site. The provided solution will be limited to the detection of cyberbullying trends in two human languages (Tagalog and English) only.