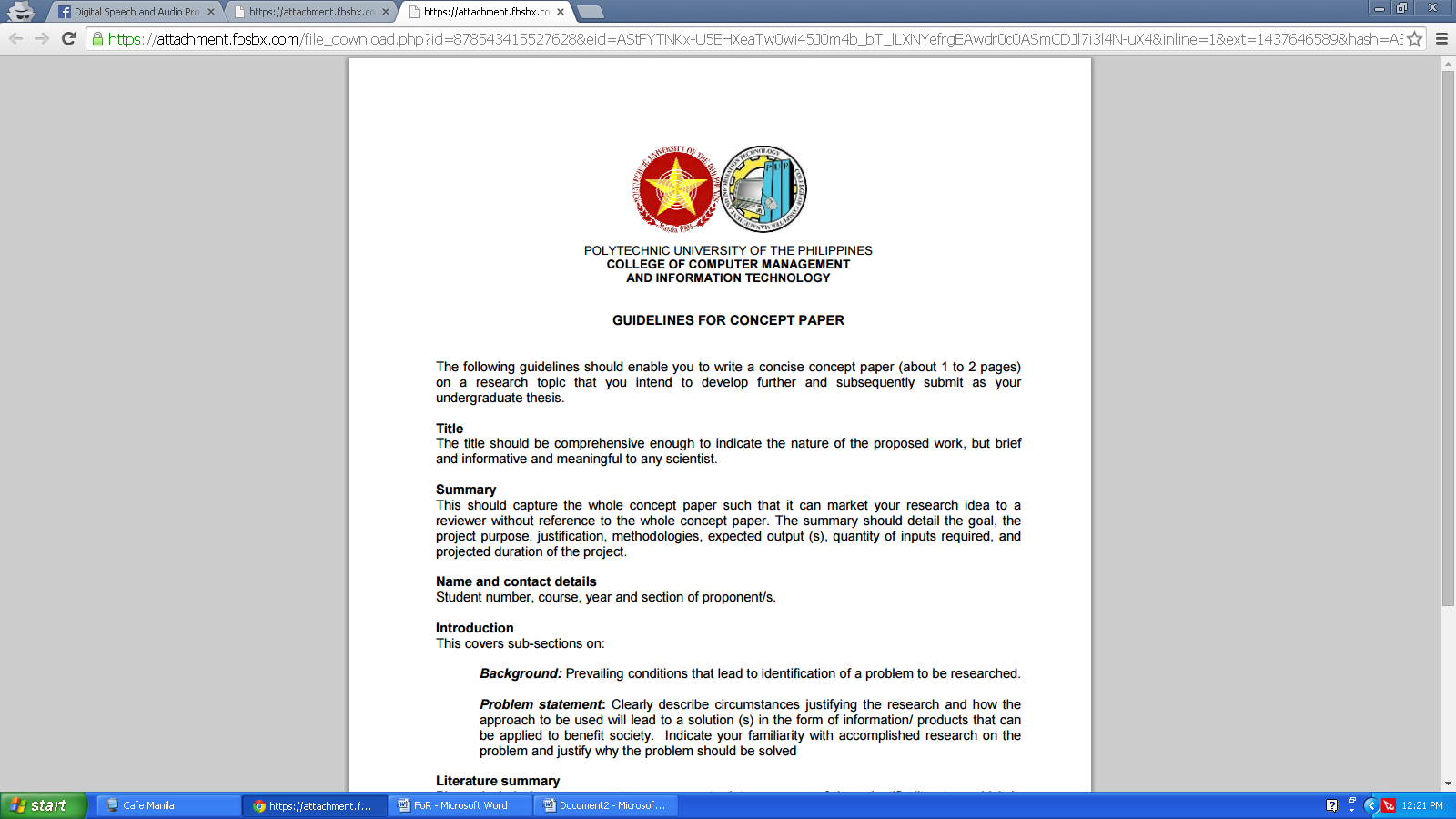
****

**POLYTECHNIC UNIVERSITY OF THE PHILIPPINES**

**COLLEGE OF COMPUTER AND INFORMATION SCIENCES**

**TITLE:**

Inappropriate Expressions Recognition using Bootstrapping as Semisupervised Learning

**SUMMARY:**

Being bullied can seriously affect a person’s physical, emotional, academic and social well-being. Many sufferers of bullying lack confidence, feel bad about themselves, have few friends and spend a lot of time alone. In this paper, the researchers will build an application that detects inappropriate expression using the bootstrapping as semi-supervised learning.

**Name and Contact Details:**

|  |  |  |  |
| --- | --- | --- | --- |
| **The Proponents (Students) BSCS 4-2** | | | |
| **Name** | Joshua Dapitan | Student No. | 2012-02498-MN-0 |
| **E-mail** | [joshuadapitan@gmail.com](mailto:joshuadapitan@gmail.com) | Phone | 09334091593 |
| **Name** | Anjanette Lasala | Student No. | 2012-03666-MN-0 |
| **E-mail** | [ahanjjj22@gmail.com](mailto:ahanjjj22@gmail.com) | Phone | 09292806420 |

**INTRODUCTION:**

Unfortunately, teasing is often part of growing up — almost every child experiences it. But it isn't always as innocuous as it seems. Words can cause pain. Teasing becomes bullying when it is repetitive or when there is a conscious intent to hurt another child. It can be verbal bullying (making threats, name-calling), psychological bullying (excluding children, spreading rumors), or physical bullying (hitting, pushing, taking a child's possessions) [Quickfind, n.d.].

In this study oﬀensive language is deﬁned as the propagation of oﬀensive messages or remarks that in some circumstances are inappropriate, exhibit a lack of respect towards certain groups of people or are just rude in general [Vandersmissen, 2012].

Classifying Inappropriate Content in English Text is important because many people who uses the internet can read inappropriate text content. Those inappropriate contents are very harmful to the readers, it can bully, harassment or etc. Like for example, anyone can bully others, they can be bullied or they may witness bullying. By creating this system, we will eventually know if it is inappropriate text content or not [Inappropriate Content, n.d.].

The algorithm to be use in this study is Bootstrapping as Semi-Supervised Learning. For Natural Language, it shows incorporating features derived from unlabeled data into a supervised model can provide substantial improvements, both in terms of reducing the error and the amount of labeled data required. Its results show that using word clusters and a new type of unlabeled data feature, mutual information statistics, can both boost performance [Liang, 2005].

**BACKGROUND:**

Bullying behavior is prevalent throughout the world and it cuts across socio-economic, racial/ethnic, and cultural lines. Researchers estimate that 20 to 30 percent of school-age children are involved in bullying incidents, as either perpetrators or victims. Bullying can begin as early as preschool and intensify during transitional stages, such as starting school in 1st grade or going into middle school. Victims of bullying are often shy and tend to be physically weaker than their peers. They may also have low self-esteem and poor social skills, which makes it hard for them to stand up for themselves. Bullies consider these children safe targets because they usually don't retaliate [QuickFind, n.d.].

In this case efficient semi-supervised learning is easy but there are five types of learning problem that have received the preponderance in machine learning. The first four are all cases of function estimation, grouped along two dimensions: whether the learning task is supervised or unsupervised, and whether the variable to be predicted is nominal or real-valued [Abney, 2008].

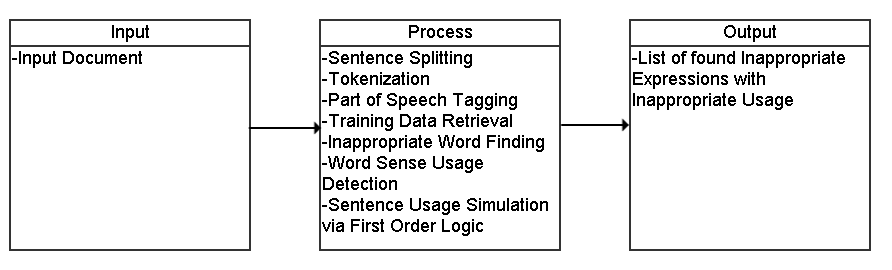
One specific open problems in semisupervised learning is the Co-Training with Linear Separators which is known that the consistency problem is NP-hard. Even if one cannot solve the problem efficiently in general, a natural question is whether one can at least weaken the independence-given-the-label assumption in a nontrivial way and still get an efficient algorithm for this class [Blum and Balcan, n.d.].

**PROBLEM STATEMENT:**

The study aims to design, develop and evaluate the System which will help to recognize Inappropriate Expressions from a document. The people who are in knowledgeable in English language are the respondents in this study. In addition to this, the researchers aim to seek answer to the following specific problems:

1. What is the performance analysis of Sentiment Analysis based on the following:
   1. Accuracy
      1. Precision
      2. Recall
      3. Specificity
      4. F-Score

**CONCEPTUAL FRAMEWORK:**

****

In this framework of the system, we see that the input is a document. We do Sentence Splitting, Tokenization, and Part of Speech Tagging as a pre-process. Then we retrieve the knowledge base of the bootstrapped data. Then we will find inappropriate words based on definition features as per the knowledge base and determine its usage. Then there will be a simulation of the sense of the word via first order logic to determine if the word is used in the inappropriate sense. The final output is the list of inappropriate expressions found in the input.

**LITERATURE SUMMARY:**

Profanity is an offensive word or offensive language [Merriam-Webster Online Dictionary, 2014]. It is also called bad language, strong language, coarse language, foul language, bad words, vulgar language, lewd language, swearing, cursing, cussing, or using expletives. This use is a subset of a language's lexicon that is generally considered to be strongly impolite, rude or offensive. It can show adebasement of someone or something, or show intense emotion. Profanity in this sense takes the form of words or verbal expressions. In its older, more literal sense, the term profanity refers to offensive words, or religious words, used in a way that shows you do not respect God or holy things, or behaviour showing similar disrespect [Longman Dictionary of Contemporary English, 2014].

An Analysis of the Pragmatic Functions of Swearing in Interpersonal Talk has shown the reason why people would choose to swear and the types of pragmatic functions which swearing carries out in everyday conversation. These functions include expressing positive emotions, including showing surprise, promoting in group membership, verbal emphasis to emphasize the speaker’s feeling about something and negative emotions, such as aggression, which threaten a person’s positive and/or negative face [Wang, 2013].

The researchers introduced the study of bullying to the NLP Community. Bullying, in both physical and cyber worlds (the latter known as cyberbullying), has been recognized as a serious national health issue among adolescents. One is being bullied or victimized when he or she is exposed repeatedly over time to negative actions on the part of others. There are wide ranges of emotions expressed in bullying traces. After manually inspecting a number of bullying traces in Twitter, our domain experts identified seven most common emotions such as anger, embarrassment, empathy, fear, pride, relief and sadness. Analyzing Social Media to Detect Cyber Bullying using Sentiment Mining found that “sentiment analysis is the task of finding the opinions of people about specific textual entities. The decision making process of people is usually affected by the opinions formed by domain authorities and the proliferation of online discussions [Bi, n.d.].

Semi-Supervised Learning for Natural Language shows incorporating features derived from unlabeled data into a supervised model can provide substantial improvements, both in terms of reducing the error and the amount of labeled data required. Its results show that using word clusters and a new type of unlabeled data feature, mutual information statistics, can both boost performance. In addition, semi-Markov models can also increase performance modestly on the named-entity recognition (NER) task but in some cases hurts performance on the Chinese word segmentation (CWS) task [Liang, 2005].

Bootstrapping requires very little in the way of modeling, assumptions or analysis, and can be applied in an automatic way to any situation, no matter how complicated. Bootstrapping fits very nicely into the data mining paradigm [Guszcza, 2005].

**PROJECT DESCRIPTION:**

The study, Inappropriate Expressions Recognition using Bootstrapping as Semi-supervised Learning is an application that identifies what is the detected inappropriate expression. The study was derived from Sentiment Analysis for Detecting Offensive Expressions, which was a Special Project in where it has worked with a littler domain, and some of the modules that will be extracted in the current project. The researchers decided to conduct this study implementing Experimental Research Method. In this method, the researcher first defines the problem to research followed by writing a research question. The performance of the system will be based on how the system detects inappropriate expressions and how the expert classifies inappropriate expressions. This study will be complete on or before March when all the testing data have been satisfied the entire documentation.

# REFERENCES

[1] QuickFind (n.d.) Bullying and Teasing: No Laughing Matter.

(<http://www.scholastic.com/parents/resources/article/social-emotional-skills/bullying-and-teasing-no-laughing-matter>)

[2] Vandersmissen, B. (2012). Automated Detection of oﬀensive language behavior on social networking sites. (http://lib.ugent.be/fulltxt/RUG01/001/887/239/RUG01-001887239\_2012\_0001\_AC.pdf)

[3] Inappropriate Content (<https://sites.google.com/site/nbushra100/home/elizabeth>)

[4] Liang, P. (2005). Semi-Supervised Learning for Natural Langauge.

(<http://cs.stanford.edu/~pliang/papers/meng-thesis.pdf>) (accessed on May 2005)

[5] Abney, S. (2008). Semisupervised Learning for Computational Linguistics.

(<https://books.google.com.ph/books?id=VCd67cGB_rAC&pg=PA4&lpg=PA4&dq=problems+in+semi+supervised+learning&source=bl&ots=Up8BGfWB62&sig=huquA4-qDJXSUpNrnkdH1T_UnL4&hl=en&sa=X&ved=0CGcQ6AEwCWoVChMIxOK6qIaSxwIVUFKOCh3oogiW#v=onepage&q=problems%20in%20semi%20supervised%20learning&f=false>)

[6] Blum, A., &Balcan, M. (n.d.) Open Problems in Efficient Semi-supervised PAC Learning.

(http://www.cs.cmu.edu/~ninamf/papers/openproblems.pdf)

[7] "Definition of Profanity", Merriam-Webster Online Dictionary, retrieved on 2014-08-31

[8] "Definition of profanity". Longman Dictionary of Contemporary English – online. (Accessed on 11 September 2014)

[9] Wang N. (2013). An Analysis of the Pragmatic Functions of Swearing in Interpersonal Talk. (https://www.griffith.edu.au/\_\_data/assets/pdf\_file/0007/589453/Na-Wang.pdf)

[10] Dr Y Bi (n.d.). Analysing Social Media to Detect Cyber Bullying using Sentiment Mining. School of Computing and Mathematics, Faculty of Computing and Engineering at the Jordanstown Campus of the University of Ulster. (<http://www.findaphd.com/search/ProjectDetails.aspx?PJID=56055>)

[11] Guszcza, J. (2005) Introduction to Boostrapping.

(<https://www.google.com.ph/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0CDEQFjABahUKEwjt_vrq0ZPHAhUDco4KHcr_CCw&url=https%3A%2F%2Fwww.casact.org%2Feducation%2Fspecsem%2Ff2005%2Fhandouts%2Fbootstrapping.ppt&ei=K-XCVa29LYPkuQTK_6PgAg&usg=AFQjCNFS7rrWAp-nzkWHWVhNeQUtfiPe2Q&sig2=wGutAiP1dGdcUnfsM7zcNg>) (accessed on September 2005)